A Machine Learning Approach to Solve the Complete Cold Start Problem in Recommendation Systems: Building the P2P Energy Trading Recommendation System with Theory of Consumption Values

Shan Shan

PhD

2022

A Machine Learning Approach to Solve the Complete Cold Start Problem in Recommendation Systems: Building the P2P Energy Trading Recommendation System with Theory of Consumption Values

A thesis submitted in partial fulfilment of the requirements of the University of Northumbria at Newcastle for the degree of

Doctor of Philosophy

Research undertaken in the faculty of

Computer and Information Sciences

May 2022

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas, and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the Faculty Ethics Committee on 03/02/2019.

I declare that the Word Count of this thesis is 43690 words.

List of publication

- Shan, S., Umar, M., & Mirza, N. (2022). Can Robo Advisors Expedite Carbon Transitions? Evidence from Automated Funds. *Technological Forecasting and Social Change*, 180, 121694.
- Shan, S., Munir, A., et al. (2021) The Role of Energy Prices and Non-Linear Fiscal Decentralization in Limiting Carbon Emissions: Tracking Environmental Sustainability, *Energy,* 234, 121243
- Shan, S., Genç, S. Y., et al. (2021). Role of Green Technology Innovation and Renewable Energy in Carbon Neutrality: A Sustainable Investigation from Turkey. *Journal of Environmental Management*, 294, 113004.
- Shan, S & li, Y. (2021), Motivations behind P2P Energy Trading: A Machine Learning Approach, International Journal of Chinese Culture and Management (In Press)
- Shan, S., Li, H., & Li, Y. (2019). *A Sentiment Analysis of Peer-to-Peer Energy Trading Topics from Twitter*. Paper presented at the 19th International Conference on Electronic Business, Newcastle Upon Tyne.

Acknowledgements

It was a difficult decision to do a second PhD before I decided to do so. I spent another three years working on my second PhD. But I would say that it was the best time of my life. During those three years, there was pain and laughter. Faced with a new epidemic, isolated at home, working, and studying at the same time, we survived these difficult times together as a family.

However, I cannot deny that I encountered considerable challenges during my research. At times I lost the courage to continue, but I would like to thank God for giving me the encouragement and support I needed to keep going.

I am sincerely grateful to my supervisors, Dr Honglei Li, Professor Longzhi Yang and Professor Zhibin Lin, who guided me through my study. They have given me a lot of help and support throughout the research process, shaping my progress.

Special thanks go to my external supervisor, Professor Zhibin Lin, who gave me great support when I lost my way in my research. He is a wise elder who has shone a light in the darkness for me.

I would also like to thank my cousin, Dr Jin Guo; my lovely friends, Miss Jun Zhu and Mr Jialou Wang, who provided me with many valuable comments and suggestions on my research; Dr Naveed Anwar, Dr Akhtar Ali and Dr Zach Lee, who provided me with expert and helpful advice on my group meetings and final oral examination.

Finally, I would like to thank my husband, Mr Yulei Li, without whose love and support I know I would not have been able to do it. Thank you to everyone who has helped me over the past three years.

Life is short, and it is hard to decide what you will not regret. However, I will be pounding the pavement for the decisions I make throughout my life. In my future academic career, I hope I can always keep my heart and keep moving forward.

Abstract

Recommendation Systems (RSs) belong to the subclass of information filtering systems, which is to determine any given user's "rating" of particular items. However, Cold Start Recommendations (CSR) pose significant challenges due to insufficient historical ratings of information. CSR is divided into Incomplete Cold Start Recommendation (ICSR) and Complete Cold Start Recommendation (CCSR) in historical ratings or user information; the former formulates recommendations to new users and new items, while the latter contributes to new systems. Various approaches deal with cold start challenges, including collaborative filtering, content-based fliting, the hybrid approach, increasing the diversity of data sources, context-awareness, and user profile construction. These approaches are commonly applied in the ICSR, though they are not as evident in the CCSR.

Few previous studies have focused on solutions to CCSR; solving CCSR is significantly important due to the popularity of the recommendation system, the increasing new system, the improved recommendation performance, and the capability of reducing user input. The Peer to Peer (P2P) energy trading recommendation system (ETRS) is an example of a new system facing the complete cold start challenge. This research compares the advantages and drawbacks of different strategies to solve CSR and selects user profile construction to solve CCSR in P2P ETRS.

First, a mixed-method data collection process was used, including semi-structured expert interviews, user-generated content analysis, and a survey to generate insights contributing to an efficient P2P energy trading user profile dataset. Second, a P2P ETRS is proposed by comparing the Decision Tree classifier and Ensemble classifier under different users' cultural backgrounds. Different cultural backgrounds may have different viewpoints on P2P energy trading. Thus, the system is evaluated with a large body of real-world data acquired from two case-study regions, one is the western dataset, and one is the Chinese dataset. The proposed recommendation system achieves high performance by delivering personalised recommendations. In addition, previous research applied to constructing the user profile dataset, demographic information, and

user personalities, neglecting the motivations behind the product selections under different cultural backgrounds. Thus, this research analysed the driving motivations of P2P energy trading with the theory of consumption values under different cultural backgrounds.

This research has two significant contributions. First, this study identified the user profile construction strategy with a machine learning approach to solve the CCSR problem in the P2P ETRS. This strategy can effectively build the original user profile dataset, reduce the user input, and evaluate the recommended model. In addition, the research adopted ensemble learning algorithms to improve the performance of the recommended model. In terms of system performance, the accuracy of the Decision Tree (DT) model achieves an 86.1% for the Chinese dataset and a 68.5% for the western data set. In addition, ensemble learning algorithms can improve the classification of both datasets, as CatBoost performed best in the ensemble model, with an accuracy rate of 94.3% on the Chinese dataset and 79.8% on the western dataset.

Second, this research created a bridge between machine learning and the traditional social science theory. A framework for a P2P energy trading recommendation system within the theory of consumption values was proposed to understand the collaborative consumption behaviour. Previous studies of consumption value theory had applied to collaborative consumption, including Airbnb and Uber. Unlike the previous research, P2P energy trading paid more attention to the feeling of uncertainty on the emotional value rather than enjoyment; That is to say, consumers believe the uncertainty of the supply of using renewable energy is more important than the consideration of enjoyment of the experience. Most notably, the essential driving values in the two datasets are different, although income can impact P2P product selection in both datasets. The feeling of uncertainty played the most crucial role in the Chinese dataset, while the western dataset focused on sustainability.

Table of Contents

LIST OF FIGURERS

LIST OF TABLES

Nomenclature

RS	Recommendation System
CSR	Cold Start Recommendation
ICSR	Incomplete Cold Start Recommendation
CCSR	Complete Cold Start Recommendation
RESs	Renewable Energy Sources
DERs	Distribution energy resources
P2P	Peer-to-Peer
IoT	Internet of Things
ICT	Information and Communication Technology
ETRS	Energy Trading Recommendation System
CS	Computer Science
IS	Information Science
CC	Collaborative Consumption
TCV	Theory of Consumption Value
SET	Social Exchange Theory
ISS	Information Systems Success Model
NLP	Natural Language Processing
DSM	Demand Side Management
LDA	latent Dirichlet allocation
CBF	Content-Based Filtering,
CF	Collaborative Filtering
HR	Hybrid Recommendation
LSTM	Long Short-Term Memory
XGBoost	Gradient Boosted Tree Ensembles
CatBoost	Category Boosting
AdaBoost	Adaptive Boosting
SVM	Support Vector Machines

CDL	Collaborative deep learning
SDAE	Stacked Denoising Autoencoder
MI	Mutual Information
mRMR.	Minimum-Redundancy Maximum-Relevance
CV	Cross Validation
GBDT	Gradient Boosted Decision Trees
CI	Computational Intelligence
AI	Artificial Intelligence
ROC	Receiver Operating Characteristics
RF	Random Forest
SMOTE	Synthetic Minority Over-sampling Technique

Chapter 1: Introduction

1.1 Research Background

1.1.1 Peer to Peer Energy Trading Recommendation System

Recommendation System (RS) belongs to the subclass of the information filtering systems, the purpose of which is to determine any given user's "rating" of an item (Resnick & Varian, 1997; Ricci et al., 2011). To generate recommendations, RS leverages users' details obtained from a registered profile, which are compared to different item reference attributes (Ricci et al., 2011). The main applications of these systems provide insights to users about items that they are likely to find useful or interesting (Schafer et al., 2001). RS plays a significant role across different industries, such as e-learning, retail, sharing economy platforms, as well as in different companies, such as Coursera, Amazon, Airbnb, and Netflix. These RSs are significant for assisting users in dealing with information overload caused by the vast amount of material available on the internet and applications, and automatically propose the most appropriate items that meet users' demands (Abdar & Yen, 2020; Camacho & Alves-Souza, 2018).

The peer to peer (P2P) energy trading recommendation system (ETRS) is a new recommendation system that recommends different energy products to customers according to their preferences. P2P ETRS has not been placed on the market yet, due to high upfront cost and physical network constraints (Ullah & Park, 2021). However, it is possible to apply to practice due to three reasons. Firstly, with the development of Distributed Energy Resources (DERs), a new distribution method allows exchange and shared local energy, which has challenged the entire power sector since its introduction. Furthermore, the emergence of innovative DERs, such as wind turbines and PV cells, allowing consumers to produce electricity, has increased the number of prosumers. Zhou et al. (2018) explain this by recognising that operators allow for a local market platform with the facilities required for consumers to reap the rewards by making it possible for prosumers to share and trade energy among themselves. This process has resulted in a new power system method with energy traded on microgrids. People selfsufficiently supply their electricity for homes, offices, and small industrial settings with Renewable Energy Sources (RESs), which can be traded and shared in local areas (Long et al., 2017). Thus, the DERs have enabled consumers to develop into prosumers. The production and supply of electricity continues to evolve and grow in the current day due to forward-thinking customers changing and replacing the traditional passive consumer market, which relies on conventional methods of energy supply. Most countries' energy supplies were controlled by large companies who had a monopoly on electricity production and distribution, meaning consumers were not involved in the supply.

However, the emergence of smart grid, Information and Communication Technology (ICT) based technology means that a large amount of data can be gathered from various sources for analysis. These sources, which include the Internet of Things (IoT), wearable devices, and sensor networks (Pieroni et al., 2018), enable bilateral communication between consumers and prosumers. This technology not only makes energy more efficient, but allows consumer data to be communicated to suppliers, allowing individual electricity usage to be assessed, coordinated, implemented, and adapted effectively (Khan et al., 2013).

Second, the increase of prosumers provides the foundation on which to create P2P energy trading suppliers (Luo et al., 2014). The prosumers who create more energy on their renewable energy sources than they can consume themselves, can sell excess to others within their local area. In other words, smaller-scale energy suppliers can also provide energy and service; as such, there has been an increase in the smaller suppliers across various sectors (Hagiu & Wright, 2015).

Moreover, new P2P energy trading has rapidly appeared worldwide, such as Piclo in the UK, Vandebron and Vattenfall Powerpeers in the Netherlands, SunContract in Slovenia, Power Ledger in Australia, and Brooklyn Microgrid in the USA (Zhong et al., 2016). Prosumers can directly trade and share energy with others via the P2P energy trading platform.

Third, the improvement of the infrastructure and technology of photovoltaic panel has led to more selection of energy products. Through P2P energy trading, renewable energy can be made accessible to individuals who have solar panel installations. This allows individuals can sell their surplus energy to their neighbours. Sellers can obtain a higher price than prices from retailer feed-in tariff; moreover, according to Aurora Energy's estimates, maintaining the power distribution infrastructure involves 41.1% of the electricity cost (IRENA, 2020). Several platforms including PowerPeers have provided customers and prosumers tracking services through which they can check where their energy is distributed and sourced (Vattenfall, 2016).

Thus, it is worth considering that the source of the energy can be selected based on its origins and where it goes, with the consumer deciding themselves where it will go. Although the energy is the same for everyone, not all individuals will need it for the same purpose, allowing diversity within the energy trading model. For example, services such as PowerPeers in the Netherlands meet consumers' demands while allowing consumers to choose where they purchase and sell energy (Havers, 2016).

Having said that, P2P ETRS are not without challenges, and P2P trading systems are still in the testing stage. Research carried out on existing trial projects reveal that poor communication between prosumers is one of the main obstacles to the successful running of P2P trading platforms (Morstyn, Farrell, et al., 2018). Prosumers have personal views and expectations on many important aspects of trading which may conflict, such as environmental or social issues, financial gain, energy security, or the level of risk they are willing to accept (Darby, 2013), which are unknown to future generations of energy traders. It is hoped that energy trading recommendations will work in a similar way to recommendations for online shoppers, predicting consumers' behaviour and desires and influencing them to buy something that accurately represents their needs (Lin et al., 2011).

The use of this system of recommendations has been studied closely in recent years across various fields such as online shopping and stock trading; nevertheless, there are no studies that analyse energy trading, particularly in the P2P trading field. Thus, there has been little progress in P2P trading beyond discussions around technological infrastructure and case studies of application. Moving forward, the next agenda regarding P2P energy is the new recommendation system for the new P2P energy trading.

1.1.2 Completed Cold Start Problem in the Recommendation System

An impotent challenge faced by any recommendation system is the Cold Start Recommendation (CSR) when recommendations are proposed (Huang et al., 2016; Schein et al., 2002). CSR means a lack of sufficient statistics; utterly unpopular items or wholly new items are introduced to the catalogue (Ricci et al., 2011). Additionally, CSR refers to a lack of credible suggestions due to a lack of ratings at the outset (Lam et al., 2008).

In other words, CSR emerges when a recommendation system is unable to create ratings for new items since ratings for new items cannot be predicted in advance. When a new item is placed on a recommendation system and no prior ratings are available for that product, or when a new user with no rating record enters the recommendation system, it is difficult to make the recommendation. According to recommendation objects, there are three types of CSR: new items, new users, and a new system (Lika et al., 2014; Zhang et al., 2010). Furthermore, Complete Cold Start Recommendation (CCSR) and Incomplete Cold Start Recommendation (ICSR) are the two major types of CSR based on whether there is historical data, (Wei et al., 2017).

Previous studies focus on solving ICSR, such as the new user recommendation (Guo et al., 2014; Karakayali et al., 2018) and new item recommendation (Aleksandrova et al., 2017; Wei et al., 2017); the new CCSR system is, however, neglected. CCSR is important nevertheless due to three reasons.

First, CCSR can provide benchmarks for new systems, especially with the increasing new system in the market. An average of 30,000 new products launch each year and with the development of web 2.0, various new platforms launched each year to match particular products (Kumar & Thakur, 2018; Tey et al., 2021). These new products and systems make the premise of the RSs. RSs can be applied in various field, such as e-commerce, retail, media, and banking (Abel et al., 2013; Goldberg et al., 1992); e-commerce has grown significantly in the internet era, leading to the sale of millions of products. Additionally, users have difficulties choosing a product from millions of options (Deshpande & Karypis, 2004). Thus, solving CCSR can effectively remove redundant information to provide users with more useful recommendations.

Second, CCSR improves the performance of the recommendation system and increases user satisfaction; without recommendations, every product has the same chance of being selected (Pu et al., 2011; Rashid et al., 2002). Furthermore, the system recommendation performance might be improved by solving CCSR. Subsequently, RSs shorten the road to make selections by offering a suitable solution.

Finally, CCSR reduces user input and user exit. Uncertainties in the information-gathering stage of a trader's decision-making process need to be removed; any user input that is unimportant to the search process, in particular, should be excluded. More parameters in the system may increase the complexity of the model, lowering P2P ERTS's performance. Moreover, reducing the user input by feature selection can increase the exit chance of the users in the early stage, as only necessary questions are asked.

Notably, however, three main research gaps exist in solving CCSR on energy trading. Firstly, existing studies focus on ICSR and rarely on CCSR; specifically, the approaches applied to solve ISCR may not be suitable to solve CCSR. To cope with the new item ICSR, collaborative filtering and content-based filtering were mostly applied to compare the similarities between different items and users (Barkan et al., 2019; Wang et al., 2018). In addition, the bulk of these studies use hybrid techniques that amalgam collaborative filtering data with Content-Based data. For example, the hybrid models Collaborative Topic Regression and Collaborative Deep Learning combine a matrix factorisation model with Latent Dirichlet allocation and stacked denoising autoencoders, respectively (Wang et al., 2015).

However, pure collaborative filtering cannot help CCSR because no user preference information is available to build any basis for recommendations. Demographic information can be useful for the CCSR, but the features of customer preference are still unclear (Ecker et al., 2018; Hackbarth & Löbbe, 2020). Thus, the approaches applied for ICSR are not always suitably applied in CCSR. Thus, in order to cope with the gap, this research adopted the user profile construction strategy to construct the energy trader user profile dataset and further to solve CCSR.

Secondly, there is a lack of studies on understanding the energy product selection behaviour under decision-making processes. Previous studies on constructing user profiles have paid more attention to demographic information (Lika et al., 2014; Safoury & Salah, 2013), social networks (Deng et al., 2014; Sahebi & Cohen, 2011), and personalities (Bai et al., 2012; Cheng & Tang, 2016; Gretzel et al., 2004). Also, previous energy recommendation research is based on the building energy recommendation model (Chwieduk, 2009; Cui et al., 2016; Yan et al., 2015), household appliance usage experiences (Aritoni & Negru, 2011; Luo et al., 2020), and neglecting motivators of product selection. Product selection motivations are important to understand users' decisions and to analyse the motivations behind their selection.

Furthermore, patterns of product selection behaviours applied in the previous P2P recommendation system may differ in the P2P ETRS, especially under different cultural backgrounds. Most P2P recommendation platforms, such as Airbnb, are based on patterns of users' behaviours and preferences, concentrating on personal matches (Abdar & Yen, 2020; Zhang et al., 2016). Grbovic and Cheng (2018) propose an embedding model created for the Airbnb marketplace; this model captures short and long-term guests' interests, resulting in excellent home recommendations.

It is worth noting that energy trading, although it is a form of Collaborative Consumption (CC) (Choi & He, 2019) when compared to the traditional power market. But it is essentially very different from other CC actives due to its own characterises. CC is when people obtain or share a resource for financial or other personal gains (Belk, 2014). CC is also typically managed online through community-based services; peers are able to share, provide, or gain access to goods and services (Hamari et al. (2016). Thus, Airbnb and Uber belong to CC under this definition.

Conversely, the particular characteristics of P2P energy trading, such as high-frequency consumption, take place daily compared to Airbnb, which had an average of 3.3 times per year in 2015 (Tussyadiah, 2016). Compared to Airbnb's low-frequency usage, there is no ownership when new users receive energy in the P2P energy trading process. The physical deterioration rate of energy is also significantly higher than room sharing or car sharing. In other words, energy is one-off consumable, whereas cars and rooms depreciate at a much lower rate. Furthermore, the marginal effect depreciation is considerably lower for attributes that guests consume (Yao et al., 2019).

In addition, recommendations might be different under different cultural backgrounds, and this is an area which lacks rigorous research. Recommendations can be made according to users' different personalities (Cheng & Tang, 2016; Gretzel et al., 2004). Concurrently, cultural factors significantly impact users' personalities (Ramírez-Esparza et al., 2006; Triandis & Suh, 2002). However, previous studies have neglected the relations between recommendations, personalities, and cultural backgrounds (Sawrikar & Katz, 2014; Schedl et al., 2018).

1.2 Research Aims and Objectives

This study aims to answer the following research questions to narrow the gaps in the research field:

RQ1: How can the complete cold start problem on P2P energy trading recommendation system be mitigated?
RQ2: What values influence energy users' consumption from the theory of consumption value perspective?

A user profile construction strategy is proposed to solve the CCSR in the P2P ETRS. Specifically, the whole process of solving the problem consists of selecting features from various sources to prepare the datasets, data pre-processing for the reduction of user input, and prosing the recommendation model. Firstly, the factors influencing prosumers' consumption of the P2P trading system is extracted through mixed methods of quantitative and qualitative studies. However, it is necessary to be noticed that due to the large number of factors which can influence prosumers' consumption, which make the data collection process is significant time consuming and costiveness. Also, the current marketing philosophy is value-oriented, and neglects to explain value from the perspective of the consumers. Thus, the theory of consumption value is applied to provide evidence of why people trade on P2P energy.

Thus, primary features that impact the selection of energy trading products were fixed in the framework of the theory of consumption values through reviewing relevant literatures. Then, features from the interview analysis and user-generated content analysis was extracted to identify the main features which can impact the product selection.

Primary data were collected through surveys internationally from China, America, Australia, the United Kingdom, Germany, the Netherlands, and China. In addition, the dataset is deemed as the user profile dataset. Finally, a machine learning approach was applied to evaluate the system performance with different algorithms. This potential new format for recommendations within the P2P energy trading system will create additional elements which will have a positive influence on the prosumers' trading experience in the P2P energy trading platform. To achieve this aim, the following research objectives are identified:

(1). Review existing studies on RSs and CSR to identify research challenges and gaps.

(2). Design and develop a survey for the user profile construction dataset under different cultural background.

(3). Evaluate the proposed recommendation model for P2P ETRS by using a variety of techniques.

(4). Reduce user input in the user profile construction.

(5). Compare different consumption values of P2P energy trading by comparing other collaborative consumption activities

(6). Understand consumers' most important perceived value when selecting energy trading platforms considering their cultural background.

1.3 Research Contributions

By examining recent research in recommendation system and P2P energy trading settings, this study makes the following two major contributions.

i. This study adapted the user profile construction strategy with a machine learning approach (a tree-based classifier) to mitigate the complete cold start problem in the P2P energy trading recommendation system. This strategy can effectively build the original user profile dataset, reduce the user input, and evaluate the recommended model. Moreover, this research adopted ensemble learning algorithms which efficiently improve the performance of the recommended model.

ii. This research created a bridge between machine learning and the traditional social science theory. A framework for a P2P energy trading recommendation system within the theory of consumption values was proposed to help to understand the consumption motivations of collaborative consumption. This research helps to understand different product selection values with other collaborative consumption actives. Compared with other P2P actives, the consumption values of P2P are different. In addition, this research helps to understand different product selection values in various cultures. Two datasets were constructed in this research, which are collected from various countries. To recognise the different consumption values played on P2P energy trading product selection in the two different datasets, the rank of feature importance is applied.

1.4 Study Outline

This study comprises nine chapters and four appendices, in addition to the introduction, listed in the following format. In Chapter 2, state of the art on recommendation system was reviewed, including content-based, collaborative, hybrid approach, and machine learning approach. Then, there is a review of the main methods of recommended systems and the cold start problem that they cannot avoid. New products, users, and systems are identified as the components susceptible to cold starts. Previous studies note the cold start from various approaches, though these approaches are still based on the collaborative filtering and content-based recommendation system. The lack of user and item data caused the basis for the cold start of the P2P ETRS, leading to difficulties using traditional collaborative filtering and content-based approaches to recommend. This also encourages the feasibilities of using a user profile construction strategy and building the user profile dataset to recommend P2P energy trading products by using a machine learning technique.

Chapter 3 outlines the possible features which may impact the P2P energy trading products. With the theory of consumption values, this chapter firstly reviews previous research on P2P energy trading. Then, it compares the differences between P2P energy trading and other P2P actives. In the final section, five sections including functional value, emotion value, epistemic value, social value, and conditional value were applied to address the selection behaviour. This chapter provides a theoretical foundation for P2P energy trading by examining previous research on the theory of consumption value on P2P platforms.

Chapter 4 outlines the approach to the study and the rationale of the proposed P2P ETRS. Expert interviews were merged with user-generated content analysis to design the survey. And then, the machine-learning approach was used to examine the recommended model.

Chapter 5, 6, and 7 provide a summary of all the subjects from the review of previous studies, expert interviews, and user-generated content analysis and used them for the survey design. Chapter 5 includes an analysis of interview responses from six participants to understand their motivations behind their selection of P2P energy products. Thematic analysis is used to extract different topics. Similarly, Chapter 6 provides an insight into the consumption values for P2P trading participation by analysing topics based on direct allocation models through user-generated contents analysis. In the chapter, the survey is designed based on the topics collected from expert interviews, user-generated content analysis, and the literature review. Chinese and western datasets were built separately to draw descriptive comparisons of the two databases in Chapter 7, such as building information, demographic information, and energy consumption information; moreover, different recommendation models are proposed for future practical scenarios.

Chapter 8 proposed the P2P ETRS by using machine learning approaches to mitigate the CCSR. To enable this, a two-part, feature-choosing method with its basis in Mutual Information (MI) in the data pre-processing part, along with a DT and ensemble-based model in the classification section, is used. Tree models were used as the classification algorithm because it provides numerous benefits, such as simplifying the analysis, meaning the recommendation selection process is transparent for decision makers. Most notably, the base DT model was compared with different ensemble models, like Random Forest (RF), XGBoost AdaBoost, and CatBoost, to determine what improvements can be made to the DT model.

Chapter 9 offers the conclusion of the study. This study suggests a decision tree-based P2P ETRS framework for decision makers, enabling energy traders to better understand how decisions are reached using a machine-learning method. The research questions are revisited in this chapter and research contributions are highlighted. Additionally, the chapter revisits the research objectives and summarises the study, providing an overview of the research limitations and proposing research insights for future studies.

Chapter 2: Recommender Systems and the Complete Cold Start Recommendation

To understand the nature of RSs and approaches to overcome CCSR, this chapter provides an overview of the well-known complexities in RS research, thus offering theoretical reinforcement for the design of an RS for the P2P energy trading system. At the outset, RS is defined, and its applications, development, and prominent techniques are examined. In turn, the CSR is contextualised, and approaches dealing with CSR were compared. Finally, the chapter examines the shallow learning algorithms including decision tree models as well as the tree family models.

2.1 Recommender Systems

2.1.1 The Concept of RS

In the early literature, RS was conceptualised as systems in which the sole inputs were user recommendations, which could thus be aggregated and delivered to those who might find them useful. The GroupLens project, undertaken at the CSCW conference in 1994 by Resnick and Riedl, was responsible for the first documented RS, and its purpose was to recommend articles of interest to researchers (Resnick et al., 1994).

In the years since the first RS was proposed and implemented, particularly with the rise of the Internet, Web 2.0, and social networking platforms, the utility of such systems has increased significantly, particularly in areas where user information-seeking behaviour is fundamental and where large volumes of information exist (e.g., e-commerce and video sharing) (Gandomi & Haider, 2015). At present, an incredibly diverse number of organisations, applications, and services rely on RS, including the playlist-generating RS used by Netflix, YouTube, and Spotify, as well as the product-based and content-based RS developed by organisations such as Amazon and Twitter, respectively. In the case of RS of this kind, they can operate using one input (e.g., video or music) or a collection of multi-platform inputs (e.g., search queries, books, or news). Pandora, Amazon, LinkedIn, and Netflix are among the organisations that utilise RS to aid their users in identifying relevant items of interest, ranging from jobs to media, thus improving the user experience and increasing revenue (Gandomi & Haider, 2015).

RS can increase the chance of purchase. As a case in point, effective RS brings users into contact with services and products aligned with their interests and needs, increasing the likelihood that impulse purchases are made. When recommendations are practical, cross-selling increases, and the available evidence suggests that the average order size also increases (Schafer et al., 2001). In the context of e-commerce, the use of RS can influence financial health, as well as the degree to which customer interactions are targeted, optimised, and positive. In terms of conversion, the way that RS improve e-commerce dialogues and customer interactions is particularly beneficial, since it increases web traffic and, for product RS and even other types of RS, purchase intention (Borràs et al., 2014). For instance, when a customer attempts to purchase a specific item in an online store, an RS can expose them to other related items at this critical point (e.g., bikes with accessories and related fitness products).

The value of RS thus also arises in relation to their ability to enhance customer loyalty through value-added relationships between the website and the customer. In settings where rival players can easily be accessed through the internet, generating customer loyalty is essential (Venkatesan & Thangadurai, 2017). For every action that a user takes on a website, the RS can receive more inputs about the user's preferences and interests, thus increasing its ability to personalise the experience; for example, the user will receive recommendations for things that are similar to those that the user has liked in the user-based CF technique. The user will receive recommendations for things that are similar to those that the user has enjoyed in the past using the item-based CF technique (Adomavicius & Tuzhilin, 2005).

2.1.2 Approach of the RS

The main techniques used in RS are content-based filtering (CBF), collaborative filtering (CF), Machine Learning (ML), and hybrid recommendation (HR). Nevertheless, the emergence of the deep learning paradigm less than a decade ago has given rise to new techniques, such as computational intelligence (CI). Ultimately, the creation of any new recommendation algorithm is based on the traditional four approaches.

First, in CBF systems, these are underpinned by the philosophy of “show more of what has already been liked”. The assumption on which CBF methods are based is that, if a user is attracted to a given item in a given category, they are also likely to be attracted to a given item in the same category, or a similar category (Wang et al., 2018). Similar features are considered in terms of how the similarity is determined (Núñez-Valdéz et al., 2012; Park et al., 2012). For example, when reviewing a Facebook video, they will receive links to similar videos on their homepage, which are selected based on CBF. In other words, CBF approaches make recommendations based on the descriptions of the items that users have rated and the descriptions of the items that should be recommended (Pazzani, 1999).

Research related to CBF are various, such as TF-IDF (Term Frequency-Inverse Document Frequency), K-Nearest Neighbour (KNN), Rocchio approach, and Decision Tree-based CBF. Taking TF-IDF as an example, it is proportional to the number of occurrences of the word in the document and inversely proportional to the number of occurrences of the word in the entire document set. In the target document, the TF-IDF of all the words in the document is calculated and compared, and the largest TF-IDF value is used to form the feature vector of the target document to represent the document. Paik and Mizugai (2010) proposed a weighted TF-IDF method on the terms folded by the title tags in RDF Site Summary contents as characteristic terms to enable users to obtain relevant RDF Site Summary contents easily and quickly.

Additionally, case-based reasoning and knowledge-based navigation systems are variants of CBF that have been discussed in previous research. For example, in the latter, the idea is to offer recommendations based on specified domain knowledge that addresses the user’s needs. Specifically, knowledge-based navigation systems match a user’s requirements to the knowledge base of that domain, and relevant, suitable, and useful products or services for the user are recommended based on their preferences (Blanco-Fernández et al., 2011). These systems provide recommendations to users based on specified domain knowledge; that is, the systems identify a user’s requirements and match it with the knowledge base regarding that domain and then recommend products which are most appropriate, relevant, and useful to the user. For instance, when an individual wants to purchase a product online, they are asked to provide their preferred requirements. Upon doing so, the system recommends the most suitable products that meet such requirements (Blanco-Fernández et al., 2011).

Second, CF methods rely on community preferences (e.g., purchase histories and user ratings). CF in information filtering was proposed by (Goldberg et al., 1992) in order to make recommendations which involve identifying correlations between users of a recommendation system. The popularity of CF stems from the fact that it is not necessary to have an explicit account of items. People-to-people relationships are considered, and the assumption is that people with comparable interests are likely to be attracted to comparable items in other areas (Liu & Lee, 2010). Thus, browsing habits, patterns, and ratings, whether implicit or explicit, can be used to gauge similarities between users. As a case in point, Facebook users see numerous suggestions about people they might know on their homepage, which are usually filtered out on the basis of several parameters (e.g., mutual friends, number of similar shared groups, number of similar liked pages, and geographical proximity) (Lee et al., 2010).

Model-based CF and memory-based CF are the two main categories of CF algorithms. Model-based CF uses a user-item database to generate an offline model and operates on reduced data, which helps to deal with scalability and sparsity issues (Krisdhamara et al., 2019; Shahabi et al., 2001), clustering (Birtolo & Ronca, 2013; Koohi & Kiani, 2016; Tsai & Hung, 2012), and dimension reduction (Adomavicius & Tuzhilin, 2005). Different from model-based CF, the memory-based CF includes the user-based and item-based two types. It makes recommendations based on user or item similarities and generates predictions for active users from the entire user-item database. User-based algorithms calculate similarities between users to recommend items based on the assumption that the interests of users will be similar in the future (Nguyen et al., 2020; Zhao & Shang, 2010). Nevertheless, the item-based algorithm looks at the similarities between items to make predictions; users are most likely to buy items that are similar to items they have previously purchased (Al-Hassan et al., 2015; Y. Zhang et al., 2014).

Third, machine learning-based intelligent RS can make highly accurate, rapid, and personalised recommendations (Batmaz et al., 2019) with the rise of novel, powerful deep learning architectures. These intelligent RSs were applied in different scenarios, such as smart heath motoring (Han et al., 2020; Manogaran et al., 2018; Manoharan, 2020), agriculture (Liakos et al., 2018; Suresh et al., 2021; Vincent et al., 2019), pattern recognition (Cai et al., 2018; Z. Huang et al., 2019), and the cold start problem of RS (Wei et al., 2017). Also, in terms of the research focus, there are two major categories that current RS research trends in applying ML methods. One is shallow learning and the other is deep learning (Goodfellow et al., 2016). These shallow designs are useful for solving simple issues and learning well-built features. For example, Ajesh et al. (2016) used the culturing and random forest to predict recommendations based on users' ratings in the online shopping scenario. In addition, deep learning as another format of ML, is also applied in RS. The applications of artificial NNs are countless, one of which is RS (Borràs et al., 2014; Yera & Martinez, 2017). As a case in point, item content features are extracted using the specific deep neural network (Vincent et al., 2010). Additionally, Collaborative Deep Learning (CDL) involves the use of deep learning techniques to improve RS, and it has been applied by incorporating a stacked denoising autoencoder (SDAE) into a straightforward latent factor-based CDL model for article and film recommendations (Wang et al., 2015).

Fourth, in contrast to CBF, CF and ML, RSs based on HR techniques are combinational systems. The underlying rationale of which features should be combined is based on whether there are clear disadvantages in one system that are accounted for by another (Woźniak et al., 2014). For example, HR systems combine the advantages of CBF and CF. For example, Netflix use a mixture of CBF and CF to recommend films to users by considering both the user’s similarity to other users and their interests (Gandomi & Haider, 2015).

Many previous studies have addressed the use of these techniques in establishing and refining RS. A summary of the findings is offered in Table 2.1.

Table 2.1: Prior studies on recommender systems, their methods, and adopted techniques

Author	Category	Application scenario	Method	Data mining technique
(Núñez-Valdéz et al., 2012)	CBF	Electronic books	Comparative analysis	Clustering (content similarity)
(Ma et al., 2015)	CBF	Friendship	Experiment	Classification (weighted similarity)
(Wang et al., 2018)	CBF	Computer science publications	Experiment	Regression (user preference)
(Bhagavatula et al., 2018)	CBF	Candidate selection	Experiment	Clustering
(Liu & Lee, 2010)	CF	Social networking	Experiment	Clustering
(Lee et al., 2010)	CF	Mobile music	Experiment	Web usage mining
(Park et al., 2012)	CF	Expert systems	Comparative analysis	Classification
(Fernández-Tobías et al., 2012)	CF	Cross-domain recommendation	Comparative analysis	Classification, clustering
(Hwangbo et al., 2018)	CF	Fashion Retail	Experiment	K-RecSys
(De Medio et al., 2020)	CF	E-Learning Platform	Experimental case study	MoodleRec approach
(Dhanda & Verma, 2016)	HR	Academic Literature	Experiment	Incremental high utility itemset (EIHI)
(Holzinger et al., 2016)	HR	Crowdsourcing	Experiment	K-nearest neighbour
(Cai et al., 2020; Katarya & Verma, 2016)	HR	Many-objective optimization	Experiments	many-objective evolutionary algorithm
(Andrawis et al., 2011)	ML	Forecasting	Experiment	Prediction (neural networks, Gaussian regression, linear models)
(Borràs et al., 2014)	ML	Tourism recommender	Comparative analysis	Classification, clustering
(Lu et al., 2015)	ML	Transfer learning	Comparative analysis	Classification (Natural language processing)
(Borràs et al., 2014; Yera & Martinez, 2017)	ML	Research topics/gap detection	Experiment	Clustering
(Z. Huang et al., 2019)	ML	Transport	Experiments	Deep neural networks
(Zihayat et al., 2019)	ML	News	Experiments	URecSYS
(Suresh et al., 2021)	ML	Agriculture	Experiments	Random Forest
(Wen, 2021)	ML	Music	Experiments	scale-invariant feature transform (SIFT), SVM and Faster-RCNN

However, compared with different studies, different approaches have their own drawbacks. CBF has two apparent drawbacks. First, CBF necessitates abundant domain knowledge because the feature representation of the items is hand-engineered to some extent (Basu et al., 1998; Goyani & Chaurasiya, 2020). As a result, the model can only be as good as the characteristics that were hand-engineered. The second limitation is that the model can only provide suggestions based on the user's current interests (Casey et al., 2008; Liu et al., 2010). In other words, the model's potential to build on the users' existing interests is limited.

The difficulties in dealing with new items is a drawback of CF (Nagarnaik & Thomas, 2015; Ramakrishnan et al., 2020). The dot product of the relevant embeddings represents the model's forecast for a given (user, item). As a result, if an item is not observed during training, the system won't be able to generate an embedding for it and will not be able to query the model.

Instead, the exploration of machine learning methods focuses on the flaws of the models. There are considerable models in machine learning and each model has its own strengths and weaknesses. Therefore, the choice of model in a recommendation system also depends on the environment. For example, deep learning is often used when recommendation systems encounter large amounts of data, but hidden weights and activations of deep neural networks are often noninterpretable according to a prominent argument against them (S. Zhang et al., 2019).

2.2 CSR in the RSs

2.2.1 The Concept of CSR

Cold Start Recommendation (CSR) is caused by lack of historical ratings of sufficient information which leads to the difficulties when making recommendations (Huang et al., 2016; Schein et al., 2002). CSR includes Complete Cold Start Recommendation (CCSR), which is a problem in which no rating records are available, and Incomplete Cold Start Recommendation (ICSR) is a problem in which only a small number of rating records are accessible for new objects or users in the system (Wei et al., 2017).

Solving the cold start problem allows for effective referrals to new users and new products. Furthermore, it can effectively solve the loss of new users and the fate of new products that are forced to be taken off the shelves because no one clicks on them (Bernardi et al., 2015; Schein et al., 2002).

2.2.2 Review on Strategies for Overcoming CSR

Various studies have identified strategies to overcome CSR (Fernández-Tobías et al., 2019; Gupta & Goel, 2018; Wei et al., 2017). There are four major strategies to note. First, hybrid algorithms may be used to overcome CSR. This approach considers the characteristics of various recommendation algorithms, based on the strengths and weaknesses of each, however CBF, CF, and ML are still the major approaches in the recommendation system (Deldjoo et al., 2019; Jingjing Li et al., 2019). Earlier research used CF and CBF (Feng et al., 2021; Lika et al., 2014), while later studies on CS integrate the ML approach with CBF and CF; for example, extracting implicit information from a user's social stream to create a user's behaviour profile by using ML techniques, such as Random Forest, which can make up for the lack of user data (Herce-Zelaya et al., 2020).

Second, the diversity of the data may be increased to overcome CSR. For example, introducing information about users' social relations with social tags and association rules (Sobhanam & Mariappan, 2013b; Viktoratos et al., 2018). Also, auxiliary information, such as attributes about users and items, was also applied to increase the diversity of the data. For user attributes, a trust network is incorporated into the raw ratings for prediction (Victor et al., 2008). Ocepek et al. (2015) combine attribute selection and local learning into the recommendation model for CS users. Not only that, with the development of deep learning, Collaborative deep learning (CDL) is a representative example that applies deep learning to recommendation systems by integrating stacked denoising autoencoder (SDAE) into a simple latent factor-based CF model for movie and article recommendations (Wang et al., 2015).

Third, context-awareness is crucial to overcome CSR. Context awareness refers to the use of additional information in the user model (e.g. location, time, weather, etc.) to improve the accuracy and real-time performance of a recommendation system (Adomavicius & Tuzhilin, 2011). In terms of the approaches, for example, contexts consist of data involving place, time, or companion. Based on the situation, it is possible to gain full or partial knowledge of contextual features. When the time principle is adopted in RS, a decay function is often utilised, whereas place-aware RS relies on wireless devices to locate users and provide appropriate guides (Fu et al., 2019). For example, Wei et al. (2017) propose to combine deep natural network and CF to extract content features of the item and use timeSVD++ to utilise temporal dynamics of user preferences. RS can solve the new item CSR by leveraging contextual data to generate effective recommendations. Companion-aware RS are directed by the concept of homophily, where Internet-based social information is leveraged to identify user clusters marked by consistent preferences and features (X. Wang et al., 2019).

Fourthly, user profiles may be constructed to overcome CSR. Elicitation of preferences entails giving the users a simple mechanism to express their preferences on objects in a dataset with minimal effort (De Amo et al., 2015). User profiles include various information. User profile construction can provide the most effective features and analyse user behaviours (Banouar & Raghay, 2016). Also, different users have different attributes, displaying different characterises and behaviours (Bergsma & Van Durme, 2013). Based on a variety of user data, a person's political inclination or ethnicity can be inferred (Pennacchiotti & Popescu, 2011).

In addition, demographic information was the eeliest approach applied to construct the user profile. Demographic information can be added to user profiles to increase the accuracy of the recommendation (Cohen et al., 2017; Gantner et al., 2010; Saveski & Mantrach, 2014). Moreover, the initial user profile may be based on the user's personality traits and use that profile to provide customised recommendations (Fernández-Tobías et al., 2016; Tkalcic & Chen, 2015). The Five factor model including extraversion, agreeableness, openness, conscientiousness, and neuroticism (McCrae & John, 1992; Wiggins, 1996) can be used as an example to make recommendations based on personal traits (Guntuku et al., 2015; Ning et al., 2019).

However, these different strategies have their own advantages and drawbacks. As displayed in the following Table 2.2.

Table 2.2 Comparison of different strategies for CSR

Strategy	**Algorithms**	**Advantages**	**Disadvantages**
Hybrid Algorisms	A mix of algorithms, based on a combination of CF, CBF and ML	Target to solve the data sparsity and ICSP problem	Requires large scale training data, and poor interpretation of the result
Diversity of the data source	Introduce user's social connections, rating comments, location information	Generate recommendations around user-centric results to alleviate cold start issues	Many parameters need to be learned, and the effectiveness of the recommendations is limited due to historical data
Context aware	Use real-time information (e.g., weather, location, mood, etc.) to generate the latest recommendations	Real-time analysis of user preferences, generating real-time recommendations.	Requires a wide range of real time data, low coverage of user referrals
Profile construction	Construct user profile, such as personal traits, demographic information	Detailed user profiling to provide more reliable user information for recommendations	Considerable labour and time costs

Moreover, Table 2.3 outlines relevant articles in a critical comparison to understand the different strategies to overcome CSR.

Table 2.3: Prior studies on CSR, their methods, and adapted techniques

Author	**Research objectives**	**Strategies**	**Dataset**	**Method**	**Data mining techniques**
(Zhang et al., 2010)	New users	Diversity Data source (Social tags)	Delicio; MovieLens	Experiments	Mathematic ranking; Inter diversity
(Sobhanam & Mariappan, 2013a)	New users	Diversity Data source (Association rules)	MovieLens	Experiments	Clustering
(Lika et al., 2014)	New users	Profile construction (Demographic information)	MovieLens	Experiments	Classification (c4.5)
(Safoury & Salah, 2013)	New users	Profile construction (Demographic information)	MovieLens	Experiments	Systematically review
(Feng et al., 2021)	New users	Hybrid	MovieLens; Film Trust; Giao	Experiments	Probabilistic matrix factorization; Bayesian personalized ranking
(Herce-Zelaya et al., 2020)	New users	Context aware (Implicit information)	Social media data	Experiments	Classification (random forest)
(J.-H. Liu et al., 2014)	New items	Hybrid	Tmall.com. Coo8.com	Experiments	K-nearest neighbour
(Aleksandrova et al., 2017)	New Items	Hybrid	MovieLens; Jester	Experiments	Matrix factorization
(Hazrati & Elahi, 2021)	New items	Hybrid	YouTube Dataset	Experiments	Neural network
(Lee et al., 2014)	New user-new items	Diversity Data source(tweets)	Twitter	Experiments	Latent Dirichlet Allocation
(Wei et al., 2017)	New items	Context aware	Netflix	Experiments	timeSVD++, neural network

Evidently, most previous studies used machine learning approaches with experiments. New users, items, and a new system are the common types of problems in the CSR. In new user CSR, the new user's preferences cannot be gauged, the new product's features cannot be identified, and the most suitable users to whom new product recommendations should be made are difficult to determine (Karakayali et al., 2018). In other words, when a new user enters an existing system, there is no user-associated data that the RS can analyse to generate recommendations. As such, the user may avoid the RS owing to the low-quality nature of its outputs. Hence, to overcome the new user CSR, collecting data from users before they enter the system, on the basis of which recommendations can be given, is worthwhile (Guo et al., 2014).

In addition, in new item CSR, products that are incorporated into an existing catalogue attract few interactions. In such a case, popularity bias occurs where unpopular items drive down the visibility of newly entered items, creating a vicious cycle wherein the RS cuts out significant subsets of the product catalogue (Bobadilla et al., 2012). Hybrid methods and context-aware are always used in the new item CSR. For example, a mixture of CF methods and CBF methods, with virtual assignment of project information files to filter the content, on the basis of which collaborative filtering is used to generate recommendations (H. Liu et al., 2014). Moreover, the dynamic scenario strategy uses an algorithm to sense the current state of new items and then selects the best item to recommend for the scenario. For example, Tang et al. (2014) proposed a meta-context strategy, whereby different scenario strategies are combined to form a pre-defined scenario on the basis of which the meta-context is divided.

However, in new system CSR (also known as CCSR), a system boots without sufficient data for the RS to use to perform effectively (Zhao et al., 2015). Previous studies on solving the CCSR elicit new user preferences by gradually probing users' response during an initial interview phase (Rashid et al., 2002; Zhou et al., 2011). As explained in Table 2.2, without any historical data, use of HR, increased data source variety, and context awareness are not able to be applied in the CCSR. Thus, profile construction is the best solution to overcome CCSR.

2.2.3 CCSR in the P2P ETRS

CCSR is significantly important for recommender systems since it takes a long time for a system from scratch to provide accurate pushes. However, it is difficult to clarify every feature when solving CCSR for a new system from scratch. Every new system is limited due to the barriers associated with evolving ex nihilo, which reflects the fact that a genuine CCSR must be endured by all systems in the journey towards application popularity. Constructing the user profile with rich information is essential to propose accurate recommendations, although there is the risk of privacy exposure.

P2P ETRS has not yet been developed specifically for renewable energy trading, which means the necessity of solving the CCSR. Despite the existence of applications such as Go Compare and Confused.com, which allow users to comparatively assess and switch utilities providers based on their energy consumption record, these platforms typically neglect the motivations that underpin consumption, and even trading, which causes the difficulties to provide personalised recommendations. In other words, the most effective information was not collected during the recommendation.

Previous P2P platforms, such as Airbnb, first estimate similarities on the basis of historical records of user activities (Zhang et al., 2010). Second, they use accessorial information, including object attributes, to filter out irrelevant objects efficiently. However, although the state of the art in RS is evolving, the key techniques have not changed, including the use of vector models to represent user preferences, as well as item expression in terms of keywords (Fernández-Tobías et al., 2012). Therefore, building keywords or attributes and the stability of preferences is critical in this research. In other words, it is necessary to build the primary energy trading profile dataset to cope with the CCSR.

A lack of historical data makes it difficult to apply CBF, CF, and Hybrid approaches. For example, conventional CF algorithms only consider the relations between items and users, which are most often represented in a U-I matrix; CF tends to perform poorly in addressing the CCSR (Wei et al., 2017). This indicates the application of the original information filtering including user profile construction and machine learning approaches are important. Thus, the recommendation system was proposed to change to find the best features to make the optimal recommendation. In other words, due to the existence of CCSR in P2P energy trading, the profile construction strategy with machine learning approach is adopted.

2.3 Machine Learning Approach

In the last two decades, with the requirement to abstract information from a large amount of data, Machine Learning (ML) as a sub-field of knowledge discovery has gained a lot of attention from both academic and industrial areas. Machine learning as an important approach was also applied in the recommender system (as shown in Table 2.1).

Most machine learning methods include shallow and deep learning (Goodfellow et al., 2016). Before the widespread use of deep learning, most Machine Learning methods used shallow architectures that only contained one or two nonlinear layers with extracted features (LeCun et al., 2015). The most common shallow learning algorithms range from decision tree-based random forest or gradient boosted tree ensembles (XGBoost) along with kernel-based support vector machines (SVM).

Deep learning can be traced back to the artificial neural network (ANN), which usually refers to MLPs with many hidden layers also called deep neural networks (DNN) (LeCun et al., 2015). Traditional shallow learning approaches rely heavily on manually engineered features designed in a specific manner for solving the task at hand. Deep learning, on the contrary, aims at automatically creating useful feature representations of the raw input data. Most deep learning models are based on neural networks, which jointly learn data representations structured in successive layers, supervised by a feedback signal (Haneczok & Piskorski, 2020).

Meanwhile, there are various types of data on the online RS circumstance. Taking e-commerce as an example, categories (e.g., gender, brand), numerical data (e.g., age, price) and multimedia data (e.g., text, images, and videos). Thus, for different data types, different approaches are applied to solve CSR. In multimedia circumstances, such as images, deep learning is most commonly used to provide recommendations as shallow learning normally cannot fit vast and complex data, and the model strongly relies on the quality of extracted features (Krizhevsky et al., 2012).

In terms of P2P ETRS, the user profile construction is an effective solution. An initial interview procedure is a simple way to get information about new user preferences (Rashid et al., 2002). During the interview, the recommendation system asks users for their thoughts on a variety of topics and creates a rough user profile.

In addition, previous studies have endorsed decision tree family models as a suitable fit for initial user interviews (Golbandi et al., 2011; Rashid et al., 2002). In a basic decision tree, each node asks users for their thoughts on a single item, and users are guided to subtrees based on whether they like, dislike, or are unsure. Finally, the recommendation list is generated using the average preference of training users within each leaf node (Sun et al., 2013). The following section introduces the decision tree and its family models.

2.3.1 Decision Tree

A Decision Tree (DT) classifier is an approach that can be used to perform multistage decision-making and is popularly applied in the RSs in different scenarios (Cho et al., 2002; Rathore & Kumar, 2017; Thiengburanathum et al., 2015). Moreover, every approach to multistage decision making aims to break down a complicated decision into a series of similar decisions to reach the desired solution (Safavian & Landgrebe, 1991). Such approaches are often employed to identify features and patterns in large databases which are crucial for discrimination and predictive modelling. These features can be interpreted intuitively, which has made them very popular for use within exploratory data analysis and predictive modelling for over twenty years (Myles et al., 2004).

ID3, C4.5, CRAT, and SLIQ are the most popular types of decision trees (Quinlan, 1986, 2014). The simplest type of DT algorithm is the ID3, although it also has several weaknesses. For example, it is not guaranteed to generate an optimal solution, it has issues with overfitting when training a dataset, and it only supports nominal variables. On the other hand, the C4.5 is an extension of ID3, which is designed to address the weaknesses of the ID3 (Quinlan, 1986). It allows for both nominal and scale variables to be used, and to address the issues with overfitting, tree-pruning is facilitated in C4.5 (e.g., confidence-based, and error-based pruning). Moreover, it allows for attributes to be missed. The C4.5 also enables information gain and gain ratio when measuring splitting, and two different splitting criteria are employed, namely: information gained and the entropy-based criterion (Witten et al., 2005). Regarding discrete data, the decision tree algorithm is highly suitable. This is because no prior knowledge is required to use it and it extracts the rules from the column data. Moreover, in comparison to other methods like neural networks, it is much easier to interpret decision-tree data. Nonetheless, it is important to note that there are drawbacks of using the decision tree algorithm, such as issues with missing data and over-partitioning of the sample space, which can ultimately cause over-fitting. Although overfitting can be avoided by pruning the decision tree, this makes the algorithm more complicated. Thus, improving the performance of the decision tree algorithm comes at a cost.

2.3.2 Ensemble Learning

Ensemble learning refers to any decision-making method that combines different inducers, especially in supervised machine learning tasks. An inducer is sometimes known as a base learner. It is an algorithm that uses several labelled examples as input to create a model (such as a classifier or regressor), in which the examples are generalised. The newly created model can then be used to make predictions for new and unlabelled examples. Any machine learning algorithm can be used as an ensemble inducer, including a neural network, decision tree, and linear regression model. Ensemble learning is developed upon the assumption that combining multiple models overcomes the errors of each single inducer. In turn, Sagi and Rokach (2018) point out that the performance of the ensemble will be much more effective than a single inducer.

The ensemble learning approach has been developed based on Kearn's proposed PAC (Probably Approximately Correct) Learning Approach, in which both a weak and strong learner are included in the system. This ultimately creates a polynomial-level learner (Kearns & Valiant, 1994). The ensemble methods discussed previously involved the use of learning algorithms that generated various classifiers and performed a weighted vote of their predictions in order to divide new data points (Dietterich, 2000).

However, in recent times, bagging and boosting methods have become increasingly more popular (Jurek et al., 2014; Rokach, 2010). Examples of bagging algorithms include Breiman (1996) Bootstrap Aggregating Algorithm and Robert's Algorithm (Schapire, 1990). On the other hand, examples of Boosting include Freund and AdaBoost (Adaptive Boost) methods (Freund & Schapire, 1997). Bagging and Boosting algorithms are both resampling methods that require data sets to be trained. However, bagging is a parallel ensemble, whilst boosting focuses on serial boosting. In other words, they both create models using different training subsets of input data (Gopika & Azhagusundari, 2014).

Ensemble learning is well applied in the RSs studies, such as recommendations on the tourisms (Nilashi et al., 2017; Wan et al., 2018); in the crop productivity (Kulkarni et al., 2018; Rajak et al., 2017); in the movies (Forouzandeh et al., 2021; Jin et al., 2005) and in the health informatics (Joshi & Alehegn, 2017; Saha et al., 2020). Ensemble learning played an important role not only applied on the RSs, but also solving the CSP challenge. For example, in Ayaki et al. (2017)’s study, they create user profiles based on product categories they visited in their access logs and predict products using Gradient Boosting Decision Tree（GBDT）, which outperforms other learning methods like SVM. Also, M. Zhang et al. (2014) present a semi-supervised ensemble learning technique to develop a model that can improve recommendation performance using context. The programme creates various (weak) prediction models using examples from various contexts, then uses the co-training strategy to allow each (weak) prediction model to learn from the others.

2.3.2.1 Random Forest

Random forests (RF) as one of the bagging approaches which are good at preventing over-fitting by combining predictors of trees so that each tree depends on the value of a random vector sampled independently and has the same distribution for all trees in the forest (Breiman, 2001). Random Forest is defined as a generic principle of classifier combination that uses L tree-structured base classifiers $\{h(X, \Theta_n), N = 1, 2, 3, \ldots L\}$, where X denotes the input data and $\{\Theta^n\}$ is a family of identical and dependent distributed random vectors (Breiman, 2001). Every Decision Tree is made by randomly selecting the data from the available data. For example, a Random Forest for each Decision Tree (as in Random Subspaces) can be built by randomly sampling a feature subset, and/or by the random sampling of a training data subset for each Decision Tree (Ali et al., 2012).

RF utilises several stochastic decision trees and combines their predictions by averaging, showing excellent performance in situations where the number of variables is much larger than the number of observations. In addition, it is versatile enough to be applied to large-scale problems, easily adapts to a variety of specific learning tasks, and returns measures of variable importance (Biau & Scornet, 2016). Generally, forest-RI is the simplest form of Random Forest, which is formed by selecting at random a small group of input variables to split on at each node. And then grow the tree using CART methodology to maximum size and do not prune (Breiman, 2001). Another approach, forest-RC, consists of defining more features by taking random linear combinations of several the input variables. That is, a feature is generated by specifying L, the number of variables to be combined. Unlike forest-RI, forest-RC first randomly selected some features and then applied a linear combination on the new features. The output of the result is decided by the following feature section and combining vote (Pretorius et al., 2016).

RF's advantages are its non-parametric nature, ability to establish the significance of variable, and high classification accuracy (Rodriguez-Galiano et al., 2012). On the other hand, as the split rules concerning classification have not been defined, RF can be regarded as a black box type classifier. Due to differences in independent identically distributed random vectors, different approaches exist. Breiman proposed forest- RI, such as the Random MultiNomial Logit (RMNL), which illustrates the Random MultiNomial Logit on a cross-sell CRM problem within the home-appliances industry (Prinzie & Van den Poel, 2008). In addition, Tripoliti et al. (2013) modified the construction and voting mechanism of RF to improve the prediction performance.

However, although the RF algorithm has been greatly improved in terms of its classification accuracy and the generalisation error, as the research progresses the problems of the RF algorithm are also exposed; for example, it cannot handle imbalanced data and the processing of continuous variables still needs to be discrete (Sun et al., 2009).

2.3.2.2 GBDT

GBDT was first developed by Friedman (2001). In gradient boosting, the negative gradient of the prior loss function is used to create a new model as it relates to the former iteration rounds. The loss function is the most significant issue to resolve in machine learning and concerns the relationship between predictions and targets; most notably, the accuracy is higher when the loss function is lower. Moreover, it can be assumed that the model follows a superior direction and changes sequentially throughout this process when the loss function appears to fall consecutively with the iteration process (Liu et al., 2017).

Given that gradient boosting is a strong and effective machine-learning method, advanced results can be obtained for many practical tasks when it is employed. It has been the predominant technique used to resolve learning problems with heterogeneous features, complex dependencies, and noisy data for a long period of time. It can also be used for activities such as web searching, recommendation systems, and weather forecasting (Caruana & Niculescu-Mizil, 2006; Wu et al., 2010); in essence, it is a process of using gradient reduction in a functional space to create an ensemble predictor. Kearns and Valiant (1994) highlight that the approach is supported by strong empirical research results that explain how strong predictors are developed and constructed by iteratively bringing together weaker models (base predictors) into a bigger, collaborated model.

The eXtreme Gradient Boosting (XGBoost) (Chen & Guestrin, 2016) and Categorical Boosting (CatBoost) (Prokhorenkova et al., 2017) are main components of the gradient boosting algorithms. As gradient-boosting methods, both algorithms recommended variants to the original algorithm to increase the training speed and enhance the generalisation capacity. XGBoost system is a scalable ensemble approach; results have shown it to be reliable and efficient in resolving machine learning issues. Additionally, a decision tree ensemble has been developed based on gradient boosting. The XGBoost system minimises loss function by expanding upon the objective function.

In addition, the intricacy of the trees is controlled by applying a variation of the loss function. This is because XGBoost only uses decision trees as base classifiers. On the other hand, to avoid changes in predictions, the CatBoost model adjusts the computation of gradients, which ultimately renders the model more accurate (Bentéjac et al., 2021). It has the capacity to manage different categorical features during training, which sets it apart from other traditional gradient boost techniques which typically focus on reducing pre-processing time and establishing effective feature-combination performance.

Moreover, when using CatBoost the entire dataset can be used for training. Prokhorenkova et al. (2017) explain that the use of target statistics to manage categorical features whilst ensuring minimum loss of information is highly effective. Specifically, CatBoost randomly permutates the dataset to identify an average label value for each example. The same category value is then placed before the one provided in the permutation. CatBoost has been found to be very effective in many disciplines, including driving style recognition (Liu et al., 2020) and predicting reference evapotranspiration in humid areas of the world (G. Huang et al., 2019).

2.4 Summary

This chapter focuses on the concept of RSs and CSR, as well as the different strategies to overcome CSR. CF, CBF, HR, and ML are the four major approaches in the RSs. Furthermore, increasing the diversity of the data source, using hybrid algorithms, context awareness, and profile construction are four strategies which may be used to overcome CSR. Specifically, the user profile construction is the best solution from scratch, although considerable labour and time costs may cause potential challenges. Previous profile constructions involve demographic information and psychology models, such as the Five factor model. The flowing chapter effectively considers the products selection preference, providing theoretical evidence to construct the user profile of a new database.

Chapter 3: Motivations of P2P Energy Trading: Theoretical Foundation

This chapter reviews the influential factors behind consumers' production selection process; it adopts the Theory of Consumption Values (TCV) proposed by Sheth et al. (1991) to review the influential factors of consumers' choice on energy trading platforms. The TCV theory aims to understand consumers' choice behaviour when they select products and services. Sheth et al. (1991) argue that the consumers' selection of a product relies on some or all of five values, namely functional value, social value, emotional value, epistemic value, and conditional value. Motivation can either originate from an evoked need, be driven entirely by external triggers, or be the result of a combination of internal and external factors. While there is a wide array of theoretical approaches are on the same, related to the buying behaviour theories. Whether from intrinsic need motivation, Maslow's hierarchy of needs theory or TCV, which map more to the dynamic relationship between different values (Sheth et al., 1991).

The concept of value is widely acknowledged as a fundamental cornerstone of marketing. The current marketing philosophy is value-oriented (Kaur et al., 2021; Kotler & Armstrong, 2013). The value refers to how customers assess the differences between all of a marketing offer's benefits and expenses in comparison to competing offers (Kotler & Armstrong, 2010). Also, the perceived value was generated from the multidimensional environment, which was impacted by price, quality, utility, and sacrifice (Sinha & DeSarbo, 1998). As one of the most classic theories in consumer behaviour, the TCV model has been widely adopted in various research contexts, such as marketing (Chen & Lin, 2019; Lin & Huang, 2012) and tourism (Choe & Kim, 2018; Kaur et al., 2021; Talwar et al., 2020). For example, Talwar et al. (2020) adopt TCV to explore factors that influence people to choose a certain online travel agency. Kaur et al. (2021) intergrade the consumption values that affect consumers' choice of the food-delivery applications. Choe and Kim (2018) argue that consumption value can impact on tourists' attitude, destination image, and behavioural intention.

Other than the relationship between consumption value and its consequences, some scholars also pay attention to the antecedents of consumption value. For example, Chen and Lin (2019) investigate what social media marketing activities influence consumers' perceived value. Jiang et al. (2019) claim that the interaction between Airbnb value facilitation and customer participation has a positive interaction on economic value; host value facilitation and customer participation, as the authors note, positively affect other dimensions of customer value, such as emotional value and epistemic value. However, before analysing the driving consumption values on P2P energy trading, it is necessary to understand the platforms of P2P energy trading as it can help to understand the previous research on this topic, the platform structural components, and types of products.

3.1 Understanding P2P Energy Trading Platforms

A review of the literature demonstrates that studies conducted from 2006 to present best indicate the period of P2P energy trading emergence, as Google trends portray a rising interest in the subject of P2P energy trading topic since 2006. Most of the studies focus on cost saving (Alam et al., 2019; Neves et al., 2020; S. Nguyen et al., 2018), community (Paudel et al., 2018; Wilkins et al., 2020), and effectiveness (Alam et al., 2017; Esmat et al., 2021). Moreover, the five major approaches adopted to study P2P energy trading are game theory (Luo et al., 2018; Paudel et al., 2018; Tushar et al., 2020), algorithms (Aitzhan & Svetinovic, 2016; Jing et al., 2020), simulation (Cui et al., 2019; S. Nguyen et al., 2018), blockchain (Cali & Fifield, 2019; Hayes et al., 2020; Thomas et al., 2019), and constrained optimisation (Alam et al., 2019; Lüth et al., 2018).

Scholars have recently started to incorporate the trading platform itself to understand the whole trading process. The most discussed components are platform setup (Talia & Trunfio, 2003; Trunfio et al., 2007), marketing mechanism (Block et al., 2008; Mengelkamp et al., 2018), price mechanism (Duan & Deconinck, 2010; Zhang et al., 2018), information mechanism (Morstyn, Farrell, et al., 2018; Pallickara & Fox, 2003; Tushar et al., 2020), and regulation (European Commission, 2015).

Integrated with the trading process mentioned above and current P2P energy trading trials (Piclo in the UK, Vandebron and Vattenfall PowerPeers in the Netherlands, SunContract in Slovenia, Power Ledger in Australia, Brooklyn Microgrid in the USA, and Noncommunity in Germany), it can be found that different approaches towards trading access are adopted by different trading trials in the setup stage. The approach used by Piclo (Open Utility, 2016, 2018) and Vandebron (Vanderbron, 2017) involves the identification of suppliers and establishing a price match for prospective customers. By contrast, SunContract (Suncontract, 2017) achieves supply-demand coordination based on its own energy pool. All schemes focus on electric energy supply, apart from Vandebron, which supplies gas as well.

However, SunContract could not gain significant insight into customers' perception of emotional connectedness to a particular supplier or into the factors influencing customers in supplier selection. Buyers can obtain energy straight from autonomous producers (e.g., farmers with wind turbines). Vandebron is similar to Piclo in that it plays the role of an energy supplier offering incentive tariffs to encourage the exchange of energy between consumers and producers. Furthermore, by giving their excess energy to Vandebron, prosumers can buy energy from the company at a more affordable cost than from other suppliers. Meanwhile, the P2P trading mechanism implemented by Power Ledger (Power Ledger, 2017), SonnenCommunity (Sonnenbatterie, 2020), and Brooklyn Microgrid (Mengelkamp et al., 2018) involved prosumers exchanging their excess energy with their neighbours. By contrast, the mechanism adopted by Vandebron, Piclo, and PowerPeers was based on energy matching, with prosumers being allowed to select local producers of renewable energy.

Regarding the market mechanism, the complete market model permitting direct P2P trading was adopted by Piclo, Vandebron, SunContract, and Brooklyn Microgrid. In this context, the platforms prioritise price matching and assisting both commercial and residential consumers in identifying the most suitable deal and price. On the other hand, a hybrid market model, integrating features from the full market and the community market, is adopted by Power Ledger and PowerPeers. This can facilitate direct trading among DERs whilst also making energy use and demand at the community level more sustainable.

Regarding the price mechanism, most schemes adopt the pay-as-bid as price model, although the Brooklyn Microgrid scheme uses a fixed rate for essential infrastructure. Acting as energy suppliers, Piclo and Vandebron implement a price match model, whereby consumers and producers are given incentive tariffs to promote energy exchange. This enables consumers to buy energy from close-by producers at a more affordable cost compared to other energy suppliers. Meanwhile, PowerPeers and SunContract allow individuals to purchase solar power plants, heat pumps, and storage units. Conversely, SonnenCommunity applies a fixed rate price for battery-derived energy, thus consumers require to assume some risk when investing in such energy. Nevertheless, SonnenCommunity assumes the wholesale price risk, thereby consumers are unaffected by that risk.

Regarding the information mechanism, Vandebron and Piclo operate as retail supplier platforms focusing on assisting suppliers in keeping their customers, as prosumers can obtain greater value from their DERs. P2P platforms afford suppliers a better understanding of their customers so that more appropriate producer contracts can be drawn up. Although it adopts an approach similar to Piclo and Vandebron, SonnenCommunity also emphasises the significance of a system of storage (Zhang et al., 2017). Thus, SonnenCommunity can be construed as a retail supplier as well as a vendor platform. Furthermore, blockchain technology is adopted by various schemes, from complete adoption in the case of SonnenCommunity and PowerPeers, to partial adoption in Brooklyn Microgrid and Power Ledger, which offer additional trading services.

Regarding the regulatory mechanism, the commercial operation of microgrid energy markets is dependent on the adoption of relevant regulations. Most countries attach significant importance to this matter, however there is no standardised regulatory framework related to renewable energy. In the UK, regulation is stringent to ensure data and privacy protection. By contrast, although regulation for data protection has been introduced in other countries in Europe, there is no stipulation about the installation of renewable energy infrastructure (e.g., PV panels and battery storage).

These different energy trading trials are similar in some respects but different in others. These various operation processes are related to P2P energy trading; they shed light on the reasons and mechanisms of development of P2P energy trading schemes which serve as the basis for the formulation of consumer choice behaviour analysis behind P2P energy trading.

However, as mentioned in the introduction, P2P energy trading is naturally a type of collaborative consumption. Thus, the consumption value on P2P energy trading may be similar as other P2P actives. The next section will review previous studies to compare the motivations between other Collaborative consumption actives and P2P energy trading.

3.2 Motivations of Collaborative Consumption

Considerable behavioural studies have focused on CC while some focus on the sharing economy. In CC and the sharing economy people work in organised networks or systems; they often participate in sharing activities such as lending, trading, renting, and swapping goods and services, money, space, and transport solutions. Some researchers consider the two to be synonymous (Benoit et al., 2017), while others believe them to be distinct concepts.

Researchers have also noted that CC is part of the sharing economy (Botsman & Rogers, 2010) and the interchangeable nature of terms such as peer-to-peer economy, CC, sharing economy, participative economy, and collaborative economy are commonly noted in previous studies (Gössling & Michael Hall, 2019). Furthermore, the sharing economy can be regarded as a normative concept in which users share underutilised assets including carpooling service BlaBlaCar (Botsman and Rogers (2010). This example best illustrates the difference between the sharing economy and CC, as consumers go to providers such as Uber and Airbnb via information platforms where they match the available goods and services with their needs.

Nevertheless, one major concern is whether it is possible to assess or transfer ownership. It has been suggested that the accessed over ownership is the most common exchange (Hamari et al. (2016). In access over ownership, users can share goods and services for a given period that can be regarded as renting, lending, or similar P2P sharing activities (Bardhi and Eckhardt (2012). On the other hand, CC includes the transfer of ownership such that ownership shifts from one user to the next in the course of swapping, donating, or buying goods, typically second-hand (Hamari et al., 2016). Moreover, in CC, the nature of exchange suggests that users are able to access resources via monetary compensation (Wittkowski et al., 2013). Studies have also shown that CC has no exchange of ownership, such as in Uber where the car's owner transfers the property rights temporarily to another actor in the transport service using a financial exchange (Benoit et al. (2017). Hence, as an economic model in the sharing economy, underused assets are redistributing through commoditising ideas related to collaboration and sharing; however, CC does not focus on ownership and instead emphasises the users' access to goods and services though information communication technology (ICT).

The question then arises as to the motivations behind CC. Various factors affect consumer behaviour such as individual responses to institutional and social norms, personal health, value for money, habits, and convenience (Clark et al. (2003). Although various studies explore CC they largely focus on behaviour and attitude. Table 3.1 presents the different motivational dimensions concerning CC. Other studies emphasise behavioural aspects by analysing individual's actions to explore feelings, human beliefs, and values as vectors for embracing the sharing economy and CC.

Table 3.1 Motivational dimensions of CC

Motivations	**Reference**
Economic benefit	(Hamari et al., 2016), (Botsman & Rogers, 2010), (Gansky, 2010), (Lamberton & Rose, 2012), (Guyader, 2018)
Sustainability	(Hamari et al., 2016), (Toni et al., 2018), (Retamal, 2019), (Amat-Lefort et al., 2020)
Culture Differentiation	(Ianole-Călin et al., 2020), (Albinsson et al., 2019), (Retamal, 2019), (Huang et al., 2021)
Social Benefit	(Oliveira et al., 2020), (Martin & Upham, 2016),

Trust	(Mittendorf, 2018), (Hofmann et al., 2017), (S. E. Lee et al., 2021), (Zarifis et al., 2019)
Perceived Risk	(Hawapi et al., 2017), (S. E. Lee et al., 2021), (Tunçel & Özkan Tektaş, 2020),
Technology	(Martin et al., 2015), (Hamari et al., 2016), (Sordi et al., 2018), (Tussyadiah, 2016), (Martin & Upham, 2016)
Community Belonging	(Bardhi & Eckhardt, 2012), (Närvänen et al., 2013), (Galbreth et al., 2012), (Neilson, 2010), (Guyader, 2018)
Reputation	(Hawapi et al., 2017), (Ter Huurne et al., 2018), (Jøsang et al., 2007), (Diekmann et al., 2014), (Mauri et al., 2018)
Ownership	(Bardhi & Eckhardt, 2012), (Botsman & Rogers, 2010), (Mittendorf, 2017), (Retamal, 2019)

Previous studies illustrate cost saving as a common incentive that drives CC (Hamari et al., 2016; Tussyadiah, 2016). Moreover, sustainability (Hamari et al., 2016; Möhlmann, 2015; Sordi et al., 2018), social (Botsman & Capelin, 2016; Botsman & Rogers, 2010), variety-seeking (Kim & Jin, 2020; Tunçel & Özkan Tektaş, 2020), enjoyment (Hamari et al., 2016; Kim & Jin, 2020) and cultural differences (Albinsson et al., 2019; Ianole-Călin et al., 2020) are also popular motivation drivers of CC.

However, P2P energy trading as a type of CC may play a different role on the TCV. CC involves intangible assets such as music, intellectual property, and space, and tangible assets such as clothes, food, and furniture. P2P energy trading is difficult to categorise as it may be deemed tangible assets as it can be controlled and measured, however it may also be deemed intangible assets as it cannot be touched. Furthermore, the consumption values impacting P2P energy trading might be different to other CC assets.

Given the different characteristics of P2P energy trading and other collaborative consumption actives, motivations influencing the recommendation system may differ. Thus, the following section reviews the TCV on P2P energy trading.

3.3 TCV on P2P Energy Trading

3.3.1 Functional Value

Functional value refers to the perceived value of a product in terms of its functionality and price (Sheth et al., 1991); it includes economic factors such as price, and product wise factors such as product quality. This concept was based on the "economical rational human" assumption in classical economic utility theory (Marshall, 2009; Stigler, 1950). Sheth et al. (1991) presumes functional value as the primary driver of the consumers' selection of products. This study identifies three main functional value which can be perceived by energy trading customers*:* economic benefits, net system benefits, and sustainability.

3.3.1.1 Economical Benefits

Economic benefits, such as price or money saving, are key determinants of choosing a product or brand. Economic utility theory presumes that individuals' actions are based on the "economic rationality" assumption, seeking to maximise utility and cost savings, or minimise transaction costs (Stigler, 1950). Several studies have shown their economic benefits (Lamberton & Rose, 2012; Möhlmann, 2015). Hamari et al. (2016) confirm the importance of price in the adoption of CC, such as P2P energy trading. In addition, Bruno and Faggini (2017) provide anecdotal evidence that CC is preferred by consumers because it allows access to the desired product at lower costs. When peer-to-peer energy trading is enabled, battery owners can buy energy from DSOs for storage at a low price and then sell it to other professional consumers at a high price(Morstyn, Teytelboym, et al., 2018). Hence, reducing prices would attract consumers to adopt certain energy products (Buczynski, 2013; Gansky, 2010; Lamberton & Rose, 2012).

3.3.1.2 Net System Benefits

One of the characterises of CC is that it is platform based due to the development of Web 2.0 (Chasin et al., 2018; Ertz et al., 2017). The information systems (IS) research community focus on investigating the effectiveness and success of such systems. The D&M Information System Success (ISS) model, proposed by DeLone and McLean

(1992), explores the relationship between net system benefits and users' satisfaction and intention to use. As such, net system benefits are some of the most essential indicators of a platforms success, which can directly impact the intention to use and users' satisfaction (DeLone & McLean, 2004; Wang & Liao, 2008). Nevertheless, it is difficult to measure and has been neglected in previous studies.

As Bock et al. (2005) argue, the net benefits in the ISS model may be influenced by many external factors; it is difficult to attribute changes in organisational performance to the knowledge management system. Thus the variable of net benefits is excluded, and the main criterion of system success is system usage. Furthermore, Chen et al. (2013) suggest that in the context of e-commerce, user satisfaction can be used as a proxy for net revenue in order to reduce the problem of overlap of variables in the ISS model. Most notably, user satisfaction and user attitudes towards the website can be used instead of net revenue to capture the success of e-commerce websites.

These studies have chosen to remove the net benefit variable from the model for a variety of reasons, which has led to a re-examination of the value and significance of this variable in the practical study of ISS models. In conjunction with the above studies that have removed the net benefit variable, they have mostly used some contextually appropriate variables as a proxy for net benefit to capture or judge the success of the information system. It can be assumed that the application of the D&M model in the real world does not mean that net income is no longer needed, but that a more specific and reasonable interpretation of this variable is required. Thus, in this research, this model summarises three major influential aspects which can influence the net system benefit: information quality, system quality, and service quality.

Information quality

System users expect timely and accurate information, as well as a consistent and easy-to-navigate interface that enables an interactive, responsive experience during transactions (Kuan et al., 2008). In terms of the information quality of energy trading, the accurate and timely calculation of energy usage reports is essential. Reliable information generated on energy trading platforms can lead to consumers' trust, which encourages them to re-use the system. For example, information quality is important to the smooth running of transactions. In the process, when the approved transaction is completed, the information from the trading process is saved in a new block that is added to the microgrid's blockchain (Yang et al., 2019).

System quality

The performance and reliability of a system can influence consumers' intention to use or re-use the system (DeLone & McLean, 1992). Swanson (1974) use the reliability of the system and the ease of terminal use to measure the quality of a system. The reliability of a system is essential for an energy trading platform as consumers expect the stable energy supply.

P2P energy trading is enabled by ICT, and thus, the use of CC by consumers may be impacted by the aspects of technology. For example, in collaborative commerce, certain significant adoption factors are ease of use, trialability, and complexity of the technology systems that enable several users to interact, transact, and collaborate with one another through an online platform (Chong et al., 2009).

In the P2P energy trading, integrating consumer participants, small-scale renewables, flexibility services, and distributed generation is difficult, especially on the blockchain platform. The operating consumer record is maintained in smart contracts that are tamper-proof, unchangeable, and transparent. The creation of such an energy trading platform can effectively offer consumers with information on energy costs and price indications(Jamil et al., 2021) .

Service quality

The quality of service of a system, as the ISS model claims, can decide consumers' intention to adopt and satisfaction level (DeLone & McLean, 1992). Energy trading platforms are services or resources accessed and provided via digital transactions (Buchanan & McMenemy, 2012). Whatever may be the causes or drivers, failure might decline trust and may result in non-recurring users and in a bad reputation. In a case of system failure, how fast and professionally the energy trading platform response is key to maintain consumers' satisfaction. The importance of service quality, together with the choice to share resources, represents unique elements in evaluating the success of the information system (DeLone & McLean, 1992).

3.3. 1.3 Sustainability

Functional value views CC as sustainable means of consuming from the perspective of utilitarian or physical attributes. From a utilitarian perspective, CC creates sustainable means to consume and benefit many people rather than a few individuals (Binninger et al., 2015). Hwang and Griffiths (2017) demonstrate that millennial consumers engage in CC because of utilitarian benefits associated with CC, like supporting members of society to acquire what they do not have partially. Sustainability is achieved because rather than owning goods and resources, sharing reduces prices associated with owning and reduces overusing resources. Möhlmann (2015) also notes utilitarian benefits such as utility and saving cost to determine the willingness and satisfaction of customers towards CC.

Sustainability consideration is a unique aspect for energy trading systems. Using natural and human resources inefficiently can harm the environment, and hence, the resource redistribution approach was developed to provide a social and economic framework to improve sustainability through effectively deploying excess resource capacity (Tussyadiah, 2016). Enhanced awareness of environmental pressure encourages people to use resources in a more efficient manner to develop a more sustainable society (Gansky, 2010). P2P energy trading platforms may decrease the negative impact on the environment by decreasing new product development and raw material consumption (Botsman & Rogers, 2010). Therefore, protecting the environment is one of the reasons why consumers adopt a certain P2P energy trading platform.

3.3.2 Social Value

According to the definition of social value by Sheth et al. (1991), the social value perceived by a consumer is achieved by the association with stereotyped demographic, socioeconomic, and cultural-ethnic groups. In other words, the choice of a specific product may result from a consumer's imagination on how other people stereotype the product after purchase. This concept is based on the theory of the leisure class proposed by Veblen and Galbraith (1973). They emphasis the influence of symbolic or conspicuous consumption value can lead to the change of consumers' behaviour. Social value of a product can be achieved by social influence, social benefit, community belongings, and subjective norms.

3.3.2.1 Social Influence

Individuals modify their attitudes, feelings, or behaviours in reaction to their society or surroundings, which is known as social influence (Varshneya et al., 2017). Social influence can be introduced through culture, nationalities, family, friends, and relatives (Moutinho, 1987; Sedera et al., 2017). Individuals' decisions are often impacted by other people. For example, if a friend suggests using P2P energy trading (WOM), but there is no trust in the platform, the importance of the person's suggestion increases in the adoption process.

Nevertheless, this social influence will not always be the same over specific distances. For example, the closer an individual is to the consumer, the more positively affected the consumer will be (Paudel et al., 2018). Additionally, Zhao and Xie (2011) investigate how peer recommendations influence decision-making. Evidently, there is a tendency that advice from individuals in the consumers' immediate social circle is accepted more than advice given by socially distant subjects. The same values are not seen in relation to time, since it is said that a person thinks more positively when the decision is associated with a distant future (Trope et al., 2007).

3.3.2.2 Social Benefit

Users of commercial collaborative systems expect both social and economic benefits to satisfy their need to expand their social relationship (Carfagna et al., 2014). Within the CC, social benefit in this study is defined as satisfaction in users' desire to become socially tied and socially connected to others. Consumers are involved in the interaction with the prosumers directly in the P2P energy trading process, which finally aids in the formation of social relationship beyond the economic and energy exchange.

Furthermore, the use of CC is influenced by social utility (Habibi et al. (2017). CC's social benefit is one of the key human drivers (Botsman & Capelin, 2016) and it is dammed as a relationship resource which will impact the exchange. According to (Cropanzano & Mitchell, 2005), the exchange of the relationship resources can result in a close interpersonal relationship, which could determine favourable and collaborative behaviours when considering social exchange theory. The rules of exchange describe how relationships develop into trusting, mutual, and loyal commitments over time. These rules accumulated from the exchange provide continuous exchange in the future. For this to happen, the parties involved need to behave according to particular 'rules of exchange', which Emerson (1976) states form a normative definition of the situation that participants adopt in the exchange relationship.

In addition, societal growth informs collaborative consumption. Hartl et al. (2016) note the motivation underpinning collaborative consumption is the need for members of society to develop. As such, collaborative consumption speeds up society's progress by ensuring that even those who need something and cannot access it can get it free or at a lower cost. For example, accommodation sharing through Airbnb and Uber for co-driving provides essential services at affordable prices, thus facilitating economic development (Wei et al., 2021).

3.3.2.3 Community Belonging

Community belonging (Närvänen et al., 2013) is a crucial driver for consumers' product selection behaviour. Sharing activities continue to necessitate a unique social infrastructure and the creation of a community in which non-anonymous agents exchange more than simply things, such as solidarity and a sense of belonging (Bardhi & Eckhardt, 2012; Celata et al., 2017). Consumers have a sense of belonging to a network of possible friends and a community of alternative travellers even if they do not know each other (Rosen et al., 2011).

Huang et al. (2021) find people engage in specific behaviours to gain a sense of identity and belonging. Moreover, McNeill and Venter (2019) identify sense of belonging and identity in a given community to influence collaborative consumption; individuals in society see it as a way of providing social support to those who cannot afford something that someone has but is not using. Thus, the well-being of society, friends, and neighbours' matter when it comes to creating social identity. Engaging in collaborative consumption improves social relations, creating a sense of identity in specific groups.

Community membership or seeking to be included in a community or group is a factor that leads to the practice of sharing or CC activities, such as energy trading (Galbreth et al., 2012; Neilson, 2010). A higher level of social trust will be established, resulting in improved transaction transparency and a reduction in fraudulent transactions in the process of energy trading. Finally, because the participants have a more direct relationship with one another, there will be a higher sense of attachment to the community (Fell, 2019; Sousa et al., 2019). For instance, people cultivate a garden to share resources, and this garden could be public or community-owned, where they can harvest fruits. This can be considered a social form of CC because consumers feel a responsibility towards the community and members (Bardhi and Eckhardt (2012).

3.3.2.4 Subjective Norms

Social value can also inform CC through subjective norms that Roos and Hahn (2017) define as perceived social pressure that significant others assert to individuals not to perform or perform a specific behaviour. Subjective norms can alter energy consumers' behaviour through impacting the social expectations of significant others such as close friends, parents or a spouse, as well as the extent to which the individual feels they need to comply with these expectations (Park, 2000). This relates strongly to the Theory of Planned Behaviour (TPB) and the Theory of Reasoned Action (TRA), both of which discuss how subjective norms and personal attitudes can affect the behaviour of an individual. Consumers' judgements about energy use are influenced by subjective norms, according to the study (Karami & Madlener, 2022; Weber, 1997).

Ajzen (1991) and Dillahunt and Malone (2015) explain subjective norms to shape sense of identity, social integration, and exchange. A sense of social identity explains why people engage in collaborative consumption. Ajzen and Driver (1992) note that TPB adds the variable of perceived behavioural control. Moreover, Ajzen and Fishbein (1977) emphasise that individuals tend to behave in a way that enables them to meet others' expectations and leads to favourable outcomes. Meanwhile, the attitude towards an event, person, object, or institution is also critical. According to Ajzen (2005), it refers to the favourable or unfavourable temperament that an individual feels, which influences their intention. However, in TRA, the outcome expectations are treated as an element of attitudes towards a behaviour, which seeks to explain the extent to which outcome expectations may be thought of as beliefs that guide behaviours (Ajzen & Fishbein, 1977).

In addition, social exchange as subjective norms influence behaviours; by engaging in collaborative consumption, people can share experiences with those who they identify with to gain reputation and social inclusion. Scaraboto (2015) claims people engage in behaviours displayed in their social group to be recognised and engage in collaborative sharing. Kim et al. (2015) demonstrates that people tend to engage in collaborative buying due to the face-saving aspect.

3.3.3 Emotional Value

The theory of consumption of value notes that the emotional value perceived by consumers can alter their selection behaviour. Emotional value infers to how consumers feel about purchasing or consuming a certain product (Sheth et al., 1991). Studies have shown emotional value to impact the use of collaborative consumption. Mayasari and Haryanto (2018) identify empathy as the key contributor to collaborative consumption. The qualitative study was also supported by de Medeiros et al. (2021) study, which found empathy, compassion, and willingness to help others access what they do not have to inform the use of collaborative consumption.

The theory of consumption of value notes that the emotional value perceived by consumers can alter their selection behaviour. Emotional value infers to how consumers feel about purchasing or consuming a certain product (Sheth et al., 1991). Both positive feelings, such as happiness and enjoyment, and negative feelings, such as anxiousness, can influence a consumer's product or brand selection. For example, negative emotions can also affect collaborative consumption. The anger and sadness of failure to help others and engage in sustainable behaviours motivated participation in collaborate consumption (Wang & Wu, 2016). For example, Hartl et al. (2016) noted that the anger due to environmental damage drives sustainable consumption through collaborative buying that ensures efficient use of resources. Furthermore, the feeling of fear that one day the environment may not sustain itself was found by Dong et al. (2018) to inform collaborative buying as a way to save resources and increase sense of pride that is associated with consuming durable and sustainable products. Emotions aroused by marketing and promotional activities may be generalised to the marketed products (Kotler, 1974; Zajonc, 1968). Many empirical studies confirm the impact of emotion on product selection (Khan & Mohsin, 2017; Liu et al., 2021). Khan and Mohsin (2017) also agree that emotional value is a determinant of green product consumer choice behaviour.

3.3.4 Epistemic Value

The perceived epistemic value of a product involves novelty seeking and variety seeking. Sheth et al. (1991) presume that the curiosity of consumers and their tendency for novelty and new experiences can influence how they select products and brands. Consumers tend to maintain a certain level of stimulation, and this tendance of novelty seeking and variety seeking can lead them to search, trial, and switch behaviours (Berlyne, 1960; Howard & Sheth, 1969). According to Minami et al. (2021), novelty influences collaborative consumption because of the aspects of newness and uniqueness of something. They conducted a study involving 400 Brazilian participants who noted engagement in collaborative consumption is informed by new way of consuming, without having to own the item at a low cost rather than buying something that is only used in short span. Supporting the former view, Mayasari and Haryanto (2018) identify trendiness and novelty as impacting collaborate consumption. Using a rational thinking approach, they identified customers as using trendiness seeking and novelty to fulfil hedonic value desires. These desires included seeking to get variety and diverse choices and selecting options that would best meet their needs.

Collaborative consumption is primarily concerned with broadening the range of commercially exchangeable goods. In P2P energy trading, the identity of traditional customers has been changed into prosumers, which satisfies novelty desire. Also, in comparison to traditional energy trading, the prosumers of P2P energy trading can offer different selections of energy sources, such as from the roof of your favourite football club, the charity you support or your neighbour.

3.3.5 Conditional Value

The last dimension of consumption value proposed by Sheth et al. (1991) is conditional value, meaning consumers' product selection can be influenced by a specific context or situation. For example, during the Christmas period, a consumer may choose a brand based on the brand's Christmas-themed design on the packaging. Different consumers from different cultural backgrounds under different situation may have different experiences (Hull, 1943). Many studies have confirmed the predictability of situational factors on consumers' behaviour (Belk, 1974; Sheth, 1973). This study identifies six major conditional values which may affect the energy trading platform selection: *demographic information; cultural difference; building information; consumers' energy consumption information; political factors.*

3.3.5.1 Demographic Information

Demographic characteristics, such as age, education, gender, and financial situation can impact the renewable energy adoptions (Rai & Robinson, 2015; Sardianou & Genoudi, 2013). Different demographic characteristics lead to the different perception and attitudes of people on products and services. For the age concern, (Warriner, 1981) discovered that because a larger proportion of the elderly live on low and fixed means, it is difficult for them to reduce consumption and make energy services a visible part of their lives. This is also supported by (Sovacool et al., 2012) who note that older people place a higher value on price stability because they have fixed incomes.

Income level can also decide the adoption of P2P energy trading. The Lawrence Berkeley National Lab recently conducted research on income patterns among US residential solar users, with a focus on low and moderate-income (LMI) consumers (Barbose et al., 2018). This report indicates that residential solar adopters have a median income that is \$32K greater than other families, and \$13K higher than owner-occupied households alone. Nevertheless, PV adoption has been heading towards more moderate-income households in recent years.

3.3.5.2 Cultural Difference

Different cultural backgrounds may lead consumers to different product selection behaviours. Individualism-collectivism is the most important feature, and it varies among countries and individuals in consumer research (Gregory et al., 2002; Han, 2017; Kacen & Lee, 2002). Previous research indicates individualistic cultures have “stronger attitude–intention and weaker subjective norm–intention relationships” whereas collectivistic cultures have “weaker attitude–intention and stronger subjective norm–intention relationships” (Cho & Lee, 2015; Singelis et al., 1995). Singelis et al. (1995) explains that personal goals are prioritised over group goals in individualist cultures, while Hui and Triandis (1986) suggest that the actions of members of individualist cultures are governed by the potential for personal gain. This explains how the extent of the influence of subjective norms varies among cultural groups.

Cultural difference aspects are social or individual influences in collaborative consumption, and the experience that one seeks when buying either individual or social experiences. Hwang and Griffiths (2017) explain cultural differences in the source of experience that individuals refer to during collaborative consumption. Unlike collectivist culture whose experience is informed by former experience of society, in individualistic societies like US, customers seek individual experience to perceive value of collaborative consumption (Amatulli et al., 2020). In an individualistic society like the US, the customer's experience with the product or service they seek informed their decision to engage in collaborative buying (Albinsson et al., 2019). Similar findings by Ianole-Călin et al. (2020), found that specific individual experiences can be shaped through the ability of collaborative consumption to create an individualistic experience, influenced by individual risk-seeking and the individual trust of a buyer or seller. However, in collectivist society, Perren and Grauerholz (2015) explain social experience as the experience of society, as people who have had a good experience when engaged in collaborative buying will then encourage others to engage in similar collaborative buying.

3.3.5.3 Building Information

The building information also can impact the renewable energy adoption because it may directly influence the price of energy consumption and installation. Building information, such as room size, has been proved to impact driving domestic PV spread by using agent based, regression, and other models. For example, Sommerfeld et al., (2017) find the number of bedrooms decides the willingness of using renewable energy by using BERT analysis. Also, the owner-occupied unit is the most influential factor in explaining the renewable energy adoption (Coffman et al., 2018).

3.3.5.4 Energy Consumption Information

Energy consumption information involves different demands of individuals. Energy consumption is impacted by occupant behaviour, weather, building envelope, house size, indoor environment, building energy and services systems, and building operation (Yan et al., 2015). It was noted that occupant behaviour significantly affects energy consumption in urban residential buildings, such as turning cooling and heating systems on and off, adjusting the thermostat, and switching the lights on and off (Yan et al., 2015). Also, the domestic energy practice in the section of the survey was concerned in relation to energy consumption behaviour, such as the weekly number of washing machine uses relates to intensity, while using a wash temperature of 30 or 40°C relates more to management (Hu et al., 2017).

3.3.5.5 Political Ideology

Political ideology can influence the perception of selection of the P2P energy trading products and the energy product selection. Political ideology refers to a set of beliefs on how to build and sustain a just and equitable society (Buel, 1972; Sen, 2009).

Liberalism and conservatism are the most discussed political ideologies in customer research (Caldwell et al., 2020; Crockett & Wallendorf, 2004; Luedicke et al., 2010). Conservatives would have an incentive to support the establishment of protective regulations for the P2P trading environment, as this would avoid disruptive changes and ensure the stability of the disrupted market (Jung et al., 2017). Libertarianism, on the other hand, is more at ease with change and uncertainty (Crockett & Wallendorf, 2004). In other words, they vary from conservatives in that they are motivated less by a desire for security and certainty.

3.4 Summary

This chapter follows the theory of consumption values framework by Sheth et al. (1991) and reviews potential functional, social, emotional, epistemic, and conditional values which may influence consumers' product selection. Apart from the common aspects, such as price and usefulness of a trading system, this chapter also reviews three factors that are unique in the energy trading industry. First, protection of the environment can drive consumers to select a certain product. Second, the building information, such as room size and number of bedrooms, can directly affect the costs of energy consumption and further influence the selection process of consumers. Finally, different political ideologies may lead to a change in product selection of consumers.

Chapter 4: Research Methodology

This chapter outlines the research approach employed in investigating the proposed P2P ETRS framework. An overall research methodology justification is developed using the support of the two-phase research design. In each study, the rationale for the research design is presented with relevant detailed research activities.

4.1 The Framework of the Proposed P2P ETRS

The Framework of the Proposed P2P ETRS adapted the user profile construction strategy to overcome the CCSR in the ETRS. It is necessary to note that asking too few questions can lead to an erroneous user profile estimates and the system failing to deliver suitable recommendations. On the contrary, asking too many questions can lead to the user exiting the initial enquiry in the decision model. In other words, the goal of an effective user profile construction is to learn as little as possible about the user's interests. It should concentrate on increasing the accuracy of the recommendations; reducing the amount of user input; and increasing the interpretability according to the theoretical foundations.

The use of the TCV theory as the theoretical foundation to construct the user profile can be justified for two main reasons. First, all five values are mutually exclusive to each other meaning that each value is independent (Sheth et al., 1991). Second, the five constructs are multidimensional and collectively exhaustive meaning that this framework has covered all aspects of consumption values (Kaur et al., 2021). The multidimensional view of this model includes both cognitive factors, such as functional value, and affective factors, such as the emotion of consumption. (Talwar et al., 2020). Woodall (2003) suggested five distinct notions of value (‘net value, "marketing value, 'derived value, 'sale value,' and 'rational value'). The value here is more about human needs and utility. In terms of P2P energy trading, except the demand of the energy and price utility, the more consumers' values reflect a concern for the environment and others, the more favourable their attitude toward sustainable behaviour becomes, and the stronger others' expectations and personal moral obligations to behave sustainably become, the more consumers will shift from unsustainable to sustainable behaviour. Thus, according to the reviews in section 3.3, under the theory of consumption value, the ability of a product to accomplish its functional, utilitarian, or physical functions is referred to as functional value. The net system benefit represents the value gained from the user's use of the system, the quality of the product and the customer service. It is also the economic and environmental value due to clean energy. In this research, the net system benefits, economic benefits, and sustainability are considered functional values. Satisfaction and feeling of security in the energy trading process determine the emotional value. At the same time, the social benefits and the social belonging that comes with energy trading; and the social influence and subjective norm that leads to the promotion and use of energy trading are all social values that come from energy trading. The epistemic value of novelty-seeking about something new and the factors under the conditional value determine the basis and prerequisites for energy trading.

The following section explains the rationale of how the proposed P2P ETRS framework to overcome the CCSR with a user profile construction strategy (see Figure 4.1). This framework includes four sub-systems developed as per data acquisition, data pre-processing, model selection, and result interpretation. This research considers the energy traders' preferences and attitudes towards P2P energy trading. Thus, the research participants are real customers with experience in P2P energy trading. Based on the consideration that cultural differences may cause differences in production selection, this research collected data from different backgrounds. Besides, P2P energy trading is discussed more in the United States of America, Canada, Australia, European countries, and China. In addition, a diverse population generally leads to precise statistical analysis because it enables more variables to be assessed simultaneously. Thus, to obtain data, the designed questionnaire was given out to users in China as well as the Western market by online survey websites namely 'Wenjuanxing' and 'Amazon Mechanical Turk (M-Turk)'. Web-based surveys represent an inexpensive, easy to conduct, and time-saving option. These questionnaires are designed based on the theoretical foundation, which can increase the interpretability of the decision.

The data was pre-processed using different data pre-processing methods, including data transformation, data cleaning, and feature selection methods. The details of data processing will be introduced in the chapter 8. This reduces the amount of user input and the risk of user exiting. Various classification algorithms were applied, such as DT, RF, AdaBoost, XGBoost, CatBoost, as all these classifiers belong to the tree-based classifiers. And then this study made a comparison between the decision tree classifier and ensemble classifier (RF and boost classifier) to investigate the different performance of different classifiers in the P2P energy trading datasets. The proposed P2P ETRS was assessed via different measurements, including the accuracy and confusion matrix and the user interface engine received the decision rules, and generated a web user interface based on the models. This is used to test the increasing the accuracy of the recommendations.

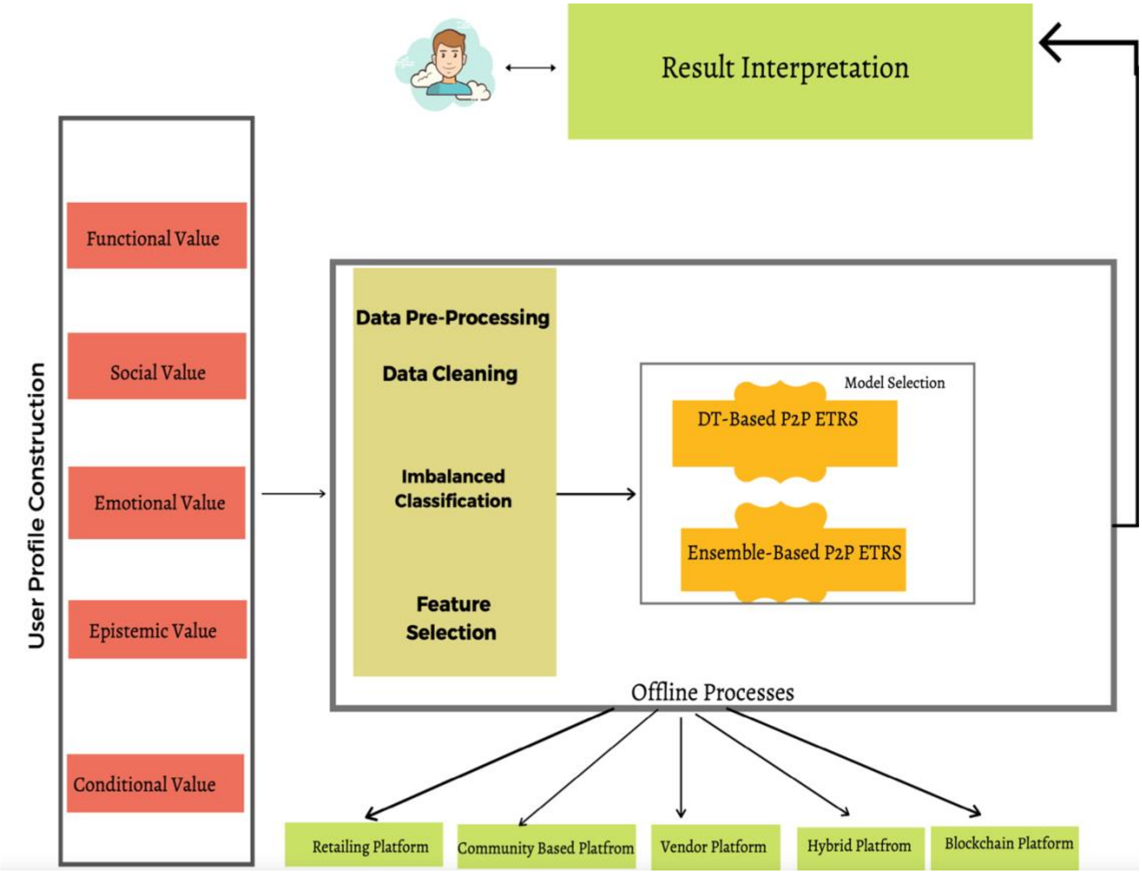


Figure 4.1 System framework of proposed P2P ETRS

4.2 Overview of the Research Approach

The study adopts a mix-method approach to design a completed recommendation system by constructing the user profile datasets. The rationale for conducting a mixed-method study is to combine the complementary strengths of qualitative and quantitative research. Typically, it is a combination of positivism and interpretivism, but also a free combination of different research methods. The qualitative research provides in-depth descriptions offering strong internal validity, while the quantitative research provides strong external validity. Neuman (2014) explains that qualitative research focuses on performance extensive studies of unique examples that happened in the natural flow of social life, while quantitative research focuses on measuring variables and evaluating hypotheses. Also, the qualitative approach is commonly used in information science because it can effectively process non-numerical data in addition to examining the data, thereby creating novel definitions, concepts, or characteristics (Maxwell, 2008).

This study took a mixed method approach to identify a convergence of the qualitative and quantitative data obtained in order to improve the credibility of the research findings (Hesse-Biber, 2010). Thus, in terms of understanding the motivations of P2P energy trading, quantitative research alone cannot provide a detailed understanding of prosumer motivations, whilst the qualitative method alone cannot validate its subjectivity and generalisability. As noted by Bryman (2016), the mixed methodology includes both qualitative and quantitative approaches. Hence, it helps address any shortcomings in either approach.

Cresswell and Plano Clark (2011) considered the combined modes of qualitative and quantitative research and divided mixed methods into four major designs, the first being a convergent parallel design in which qualitative and quantitative research are weighed simultaneously and equally. The second method comprises an exploratory sequential design in which qualitative data is gathered and assessed prior to quantitative research. In the third method, there is an explanatory sequential design in which quantitative data and then qualitative data is gathered, whilst the fourth method is the embedded design that is applied if the researcher finds either qualitative or quantitative research to be insufficient and needs additional data to add value.

The current study adopted the exploratory sequential design approach, which starts with a qualitative phase that includes data collection and analysis. Various qualitative methods can be employed, such as interviews, case studies or thematic analysis in the qualitative phase. Results generated from the qualitative phase are used to produce or inform the subsequent quantitative phase, which can involve a survey or other forms of quantitative data collection (Cresswell & Plano Clark, 2011).

Exploratory sequential mixed methods design was used to conduct a survey to capture features of P2P energy trading consumption behaviours in multicultural scenarios. The rationale for this selection is that the central issues with P2P ETRS is the CCSR. An energy trader's selection of their desired platform is significantly affected by their trading preference. Offering recommendations from the outset is challenging a clear understanding of the preferred platform of a trader is difficult to determine after the trade has begun. Also, CCSR exists in P2P energy trading which means the data has to be collected in the form of a survey as opposed to secondary data.

In addition, the dearth of data pertains to both items and users. Hence it is necessary to examine considerable material to locate the principal determinants of energy trading, after which the P2P ETRS is designed for the users. The mixed methodology examines the factors shaping energy trading and offers precise, tailored trading recommendations. The mix-method approach includes an expert interview, User-Generated Contents analysis, survey, and machine learning, in addition to the entire approach flow, as indicated in Figure 4.2.

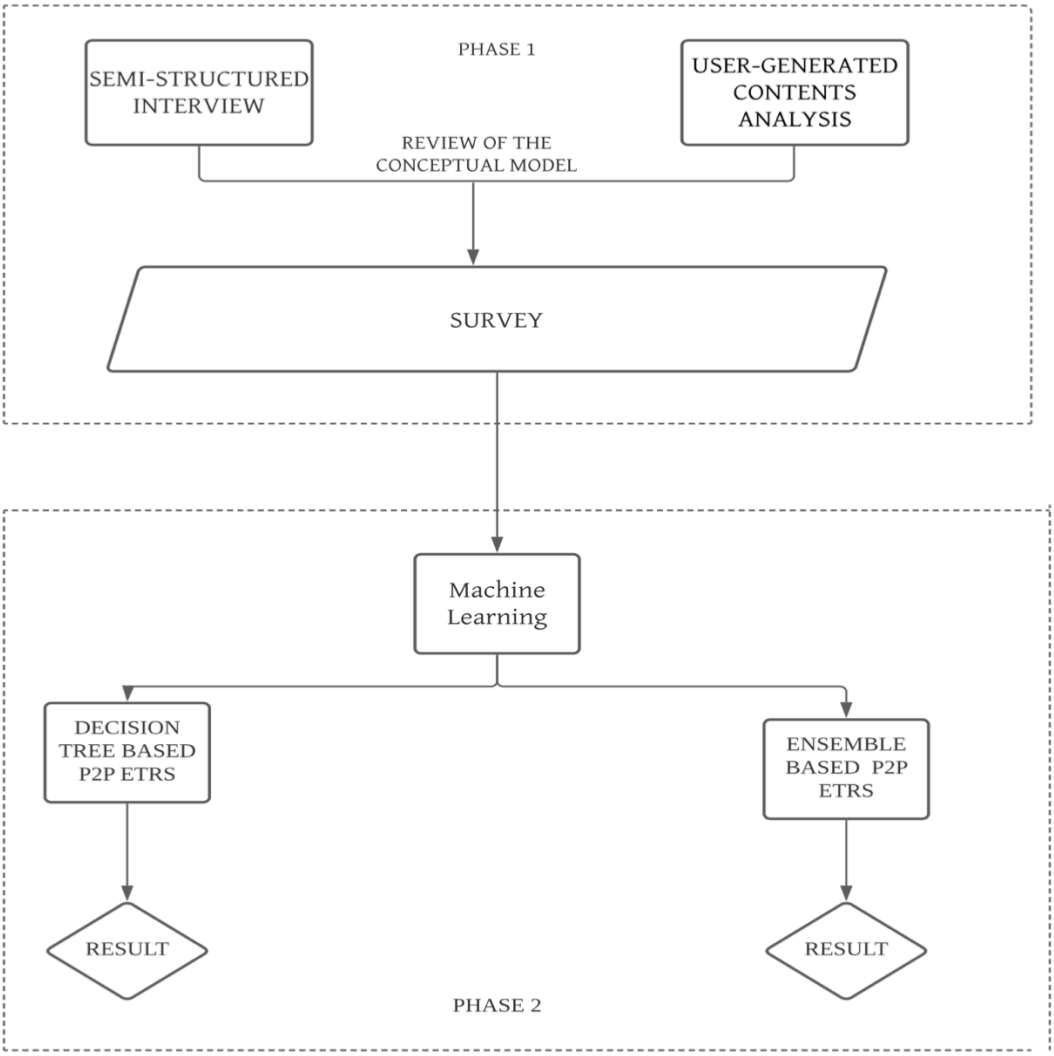


Figure 4.2 The research process flow of proposed P2P ETRS framework

A mixed method was applied in the first phase of the study. An essential qualitative phase was designed to construct the P2P energy trading dataset and the data collected by the survey. However, to capture extra energy users' trading motivations, this study applied expert interviews and user-generated content analysis. Six interviewees from different components of the p2p energy trading platforms were involved in the expert interview. They were contacted by mobile phone after getting their permission. All the transcripts are recorded and transferred into both Chinese and English. While the data from user-generated contents are extracted from Twitter and Vandebron to cope with different customer channels. Thematic analysis is applied to analyse the interviews, while a machine learning topic modelling is used to analyse the user-generated content.

The findings were integrated into the design of the survey to measure the trading intention features in different cultural groups. This provides the original evidence for the establishment of the P2P ETRS dataset.

The second phase is the quantitative phase including the machine learning approach. Two experiments were employed, the objectives being to examine the performance of the decision tree-based classifier and to determine how the ensemble-based classifier can render the P2P ETRS classification more accurate. The two experiments located appropriate features and optimal models from the different algorithms, as well as dealing with feature selections and the imbalance of data classification. Furthermore, the ensemble-based classifier compares the selected classifier results to determine whether elevated classification accuracy is achievable.

4.3 Data Acquisition

Using user profile construction strategy to solve the CCSR requires a dynamic dataset to meet the long term and short-term requirements. Previous studies mostly use psychology models, such as the Five factor model (Devaraj et al., 2008; Kosinski et al., 2014; Souri et al., 2018) or the ten item personality inventory (T. T. Nguyen et al., 2018; Oshio et al., 2013) to measure personalities and further construct user profiles. This study used TCV model to construct the user profile dataset and further mitigate the CCSR in the P2P ETRS. Also, there are some significant areas of information around the trading intention to use P2P energy trading systems that would be better retrieved through tools such as interviews and the use of social media. Previous studies carried out using purely quantitative or qualitative approaches do not paint the most accurate, rounded picture of public opinion and intention towards P2P platforms.

Figure 4.3 demonstrates the data acquisition structure. There are three steps in the dataset construction. The researchers extracted factors from the expert interviews, and user-generated content analysis, after which they combined the factors with the literature review and summarised all of the factors pertinent to the questionnaire items. As Fowler Jr and Fowler (1995) mentioned, prior to developing the questionnaire, the objectives and answers required to satisfy them must be defined. The objectives are to analyse different users' consumption behaviours and identifying the motivations behind P2P energy trading. Consequently, the researcher designed the survey questions so they would provide supporting evidence for the factors impacting trading and final purchase choices.

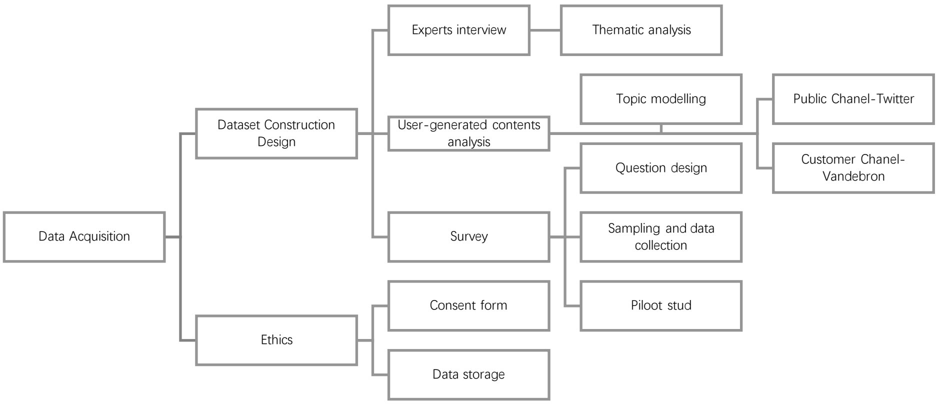


Figure 4.3 The Structure of Data Acquisition

4.4 Ethical Consideration

Since this study includes human participants there are certain ethical considerations that must be addressed. It is crucial for researchers to obtain formal ethical approval from the relevant ethical committee prior to obtaining participants for their study. As noted by Collis and Hussey (2013), a study concerning human participants must address several ethical issues prior to gathering data. As this study aims to assess consumer perceptions and behavioural intention regarding P2P energy trading, the interviews were conducted first, followed by a survey questionnaire. This process facilitated the gathering of data from a large sample. The Computer and Information Science Department's Ethics Committee granted ethical approval for this study with approval number 14207.

Principal ethical concerns in this study included the need for anonymity and confidentiality. Enabling participants to be anonymous leads to higher response rates, ensures accuracy, and increases honesty (Collis & Hussey, 2013). Thus, in the expert interview, respondents were allocated coded numbers as substitutes for their names, with M indicating male and F indicating female, which ensured anonymity.

Furthermore, the gathered data was utilised only for academic purposes, thereby ensuring that participant information could not be traced back to them. All collected data was saved on a password-protected laptop to which only the researcher has access. Hard copy versions of the questionnaire data were also locked in the researcher's personal cabinet. After completing the study, the hard copy records can be sent to the offsite storage facility of the university and the electronic materials can be archived on the information science department computer.

Participants were requested to sign an information and consent form prior to the interview and questionnaire. The consent form outlined the research goals, the procedure involved, and the methods utilised to safeguard personal data. Participants were required to agree to participate in both studies after carefully reading the informed consent form (Appendix D). They were informed of their right to withdraw from participation at any stage without the need to offer any explanation.

Chapter 5: Phase 1- Expert Interviews

This chapter covers the experts' interviews regarding the five components of P2P energy trading platform in section 3.1. The main task was to investigate if any new themes emerged from the interviews after comparing the factors that TCV impact on CC in the literature review. Based on the knowledge developed and the topics addressed, this study formulated an expert interview to ascertain if any new topics were generated. The interview has two primary aims. Firstly, to know the experts' perceptions of P2P energy trading platforms and if there are challenges and opportunities in the trading process. This is accomplished by having the interviewees respond to a series of pre-formulated questions. Secondly, to sort the narrative information into different content categories which disclose similarities and distinctive perspectives among the interviewees, followed by the consideration of capsulised core themes and subthemes regarding the acceptance of P2P energy trading. This procedure is critically important as it not only offers detailed information regarding experts' perceptions towards P2P energy trading platforms, but it also provides some clues for developing a questionnaire in the subsequent quantitative study. Thus, this chapter is divided into three sections including research methodology, data collection, and findings with discussion.

5.1 Interviews

Interviews are used to determine how respondents act, feel, and think and are efficient data collection tools (Collis & Hussey, 2013; Crowther & Lancaster, 2012). Expert interviews are widely used in the social and political sciences (Van den Bulck et al., 2019). They are a qualitative empirical research approach that are used to investigate expert knowledge (Meuser & Nagel, 2009). Experts who have the capability for knowledge development, rather than the possession of a high hierarchy and social status, were involved in the expert interview (Desmond, 2004). Expert interviews are a good way to acquire data, especially when researchers are looking into a new, emergent, or under-studied topic (Littig & Pöchhacker, 2014; Y. Wang et al., 2019).

Expert interviews were chosen to learn more about how P2P energy trading practitioners navigate and make sense of the new P2P ETRS landscape. Experts in a given problem area may have knowledge of a system's causal structures and are likely to have reasonably well-grounded assessments of the current state of affairs in the areas of concern (Scapolo & Miles, 2006).

The expert interview takes a semi-structured style focusing on understanding the perspectives of participants concerning P2P energy trading and enabling them to express opinions in their own words. As stated by Crowther and Lancaster (2012), semi-structured interviews tend to be flexible. Moreover, they enable researchers to guide respondents towards certain subjects. Researchers can also ask probing questions and request additional information or clarification. In semi-structured interviews, the interviewer starts with a laddering question to stimulate a pertinent narration. After the interviewee had finished the narrative episode, the interviewer progresses to general, specific, and ad hoc questions. Through general explorations, researchers secure data pertaining to areas of potential research interest that have not been mentioned by the interviewee. Specific explorations refer directly to the interviewee's accounts to gain further insights and opinions. The questions and statements are based on knowledge that was gained beforehand or that emerges during the narrative.

5.2 Data Collection

In qualitative research, sampling is conducted to obtain detailed knowledge of a phenomenon by assuming that the individual is a typical representation of the group/ This can offer insights into the group therefore sample size or sampling strategies are not crucial for sample selection (Kumar, 2018). In qualitative research, the sample size is determined in a subjective manner. Thus, no predetermined sample size is utilised in qualitative research, as noted by Kumar (2018). After the data reaches the point of saturation, it is not possible to collect any new information from other respondents. Hence, for the expert interviews, a qualitative method is used where in-depth interviews are conducted to gather information about consumer perceptions and their intention to utilise P2P energy trading. A judgement based sampling method was also utilised to select "information-rich" respondents with relevant professional experience, instead of random sampling (Hoepfl, 1997). The research sample comprised six participants of different nationalities, social statuses, age groups, educational backgrounds, and incomes. From these interviews, significant information concerning P2P energy trading motivations was collected to help identify key abstract themes or constructs relevant to the subsequent study.

Six interviews were conducted with four male and two female participants from 7th July to 18th September in 2019, with the interview duration varying from 30 to 50 minutes. To ensure the anonymity of the participants, the capital letters M and F were used to represent gender (e.g., M1 is the first male participant and F2 is the second female participant). All interviewees are from different backgrounds and interviews were conducted virtually using online Zoom meetings. Table 5.1 presents the six interviewees who participated in this study. The interviewees were selected because they construct the major components of the trading process. They are confident answering questions in relation to the trading procedure and regulations. Two interviewees were line managers from the Shanghai energy conservation and reduction centre (NGO), one was a researcher from Neimenggu's Electric Power Research Institute, one was a manager from China's state grid corporation, and the remaining two were operations managers from a renewable trading platform in the UK.

Table 5.1: Statistical Table Presenting Interview Respondents

Interview Sector	Energy NGO	P2P Energy Trading Platform Staff	Government Officer	Researcher	Total
Respondents	2	2	1	1	6

Combining an open narrative with a more structured interview section enables the researcher to remain receptive towards the information, since this approach encourages interviewees to describe individual perspectives, clarify previous statements, and revise misguided assumptions of the interviewer. Furthermore, these interviews were guided by researcher questions and expressed the theoretical propositions developed in the analytical framework. Interview questions are designed based on the most discussed components of P2P energy trading platform (section 1.1.2) including set up, market mechanism, information system, price mechanism, and regulation. This ensures that the participants were focused on these subjects (see Appendix A (The Interview Schedule Chinese/English), including the Chinese section and European section.

Centring on the above five categories, laddering allows the researcher to explore the participants' understanding of a particular issue regarding P2P energy trading platforms. A laddering interview is frequently considered to be part of the Repertory Grid method in IS (Tan & Hunter, 2002). The objective of laddering interviews is to capture both the content and the structure of these personal constructs, as these are both the antecedents and outcomes of people's attempts to order their lived experiences (Schultze & Avital, 2011). In its structurally predefined form, this interview technique can be effectively employed by researchers (Bannister & Fransella, 2019). Therefore, the researcher arranged laddered questions in an order that starts with the least invasive questions and proceeds to the most invasive questions (Price, 2002). The laddering question for every process is as follows and, taking energy set up as an example:

i. Can you describe how the current P2P energy set up works?
ii. What do you think are the main challenges in the current energy setup?
iii. Are there any challenges and opportunities for the future of P2P energy trading?

Onwuegbuzie and Leech (2007) stated that, when an interview is conducted, it is not merely a neutral exchange of questions and answers, but it involves conscious and unconscious motives, feelings, and desires that cannot be avoided. Thus, awareness of interview bias was necessary before conducting the interviews. In this study, the interviewees were deemed to be experienced in constructing in-depth observations. Despite the interviews being loosely based on a thematic framework for examination, they were conducted in an open-ended manner, as per the semi-structured interview method, to secure a detailed understanding of interviewee experiences. In addition, the interviews were performed in an interactive-relational way, so the researcher was aware of their position in this study and conveyed the interview aims to the respondents. The features identified in Chapter 3 were also added to the interviews and the researcher's personal views were avoided as much as possible.

Interviews are often recorded by taking notes because this enables researchers to begin assessing and interpreting earlier in the research process (Motowidlo et al., 1992). Note-taking enables the researcher to conduct partial analysis and focus on the interviewee's responses. However, this method is limited to non-verbal contact because the researcher may focus more on taking notes than having natural interactions (Motowidlo et al., 1992). Therefore, in the present study, the researcher took notes to record the interviews by focusing on sentences and keywords related to P2P energy trading and observing these throughout the conversation.

5.3 Data Analysis

Template analysis is one way of conducting thematic analysis which allows researchers to use hierarchical coding to analyse textual data, and is flexible enough to adapt to specific studies (King, 1998). Template analysis means that researchers can examine the data and list topics while grouping them. Thereafter, thematic analysis can be conducted to explore the content of the transcription of interviews, thereby defining themes within the data and organising those themes into some type of structure for better interpretation (Brooks et al., 2015). Repeated words, phrases, or narratives can be imposed by the researcher, as derived from the initial narrative data. In other words, the purpose of template analysis is to create a coding template based on a subset of the data, which is then applied to further data, revised, and refined in a flexible style and format (King, 1998).

The researcher followed (Brooks et al., 2015) in their template analysis to understand expert thoughts regarding the P2P energy trading project. Firstly, an initial narrative was generated to describe the participants' main thoughts regarding different operation process of P2P energy trading. Each of the original interviews lasted for a period of 25 minutes. The necessary transcripts were extracted and translated using both English and Chinese. The initial transcription of interviews was pre-identified by coding via the highlighting of the repeated words, phrases, or narratives which reflected their perceptions of different values (functional value, social value, emotional value, conditional value, and epistemic value).

For the second step, the target was to organise the new emerging themes into different clusters with an initial version of the coding template to be developed afterward. The emerging themes were organised into different clusters based on the frequency described by the interviewees (e.g., supply issue, transfer cost, product quality, and environment issues). This was then applied to additional data to ascertain whether new themes were generated. Then, the pre-designed themes were modified to ensure that themes could cover information within the new data. For example, whether the feeling of security could be categorised as an emotional value or whether the transfer cost could be categorised into functional value. Finally, the core themes and sub-themes were finalised to ensure no extra information needed to be coded.

5.4 Findings and Discussion

The interviewees were aged between 32 and 55 years, and four out of six participants were within the age range of 36 to 45 years. To ensure a sufficient level of diversity, all six participants came from different departments related to the P2P energy trading platforms, including two Energy NGO (33.3%), two P2P platform staff (33.3%), one government officer (16.7%), and one researcher (16.7%).

The hierarchical coding structure was formed with 10 core themes, and the frequency of the sub-themes are demonstrated in Table 5.2. These 10 core themes are summarized from the previous literature review under different theory of consumption value framework, as per see in chapter 3. By doing so, the key categories are expanded to 10 key core themes and the following section will feature a detailed analysis.

Table 5.2 Generating Initial Codes

Sub-themes	Mentioned times	Core Themes	Categories
Policy	16	● Political ideology	A: Conditional Value
government	13	● Culture differences	
order	6		
Individual	7		
Culture	4		
China	5		
UK	15		
US	4		
Different	6		
decentralized	6		
Agriculture	4		
Organization	5		
Control	6		
Centralized	8		
EPC	5		
Ideology	3		
Trading object	25	● Social influence	B: Social Value
Neighbour	7	● Subjective norms	
Friends	15	● Community	
Internecion	15		
Network	6		
Community	5		
Household	4		
Family	2		
Safe	4		
Achievement	5	● Felling of security	C: Emotional Value
Worried	6		
Satisfied	15		
Enough	10		
data	9		
Cryptocurrencies	17		
Trust	5		
Privacy	4		
Bitcoin	5		
e-coin	10		
Ethereum	4		
Supplement	7		
Allowance	14		
Integrated	15	● Net system benefits	D: Functional Value
Stable	6	● Economic benefit	
System	18	● Sustainability	
Efficiency	5		
Transaction	20		
Control	5		
Batteries	16		
Storage	8		
Weather	19		
Cost	9		
Install	9		
Promotion	4		
Cheaper	11		
Allowance	18		
Supplement	12		
Renew	35		
photovoltaic	28		
Environment	12		
Solar	6		
Source	7		
Natural	7		
Recycle	3		
System transfer	4		
Experience	11	● Novelty seeking	E: Epistemic Value
Try	5		
Want	8		
Technology	12		

5.4.1 Functional Value

i. Net system benefits

The three dimensions of net system benefits (including information quality, product quality, and system quality) are reflected within Table 52. All of the interviewees mentioned the usefulness of the platform, although it was discussed in different aspects. For the system quality, users are more concerned about the status of the platform and how well it works. As M4 explained:

‘Many users respond that they are not sure how to exit the system in the future and if it is easy to exit, or to switch to the original system’.

The stability of energy supply is critical in energy trading systems. The stable energy supply can impact the energy price (Wang & Sun, 2017) and battery storage is acknowledged as critical to increasing customer flexibility, maximising trading revenue, and maintaining market balance (Lüth et al., 2018).

‘*Many customers are very afraid that if energy is not connected to the mains network, they will have no electricity for themselves if the weather is bad’*.

Also, it is important to provide the stable supply so energy storage impacts product selection. For example, as M1 mentioned,

‘Customers consider the storage of electrical energy when installing solar light panels, and high-capacity batteries are a better seller’.

As well as M1 considering electricity storage and weather impact, he states that customers choose the P2P energy trading platform as:

‘Low energy storage can cause the increase of the energy price, and so, they do not like to trade energy when batteries are low and bad weather’.

When the state of charge was low, there was a significant drop in desire to sell self-generated electricity, indicating that state of charge has a significant impact on trading decisions (Ecker et al., 2018; Hahnel et al., 2020).

For information quality, Kuan et al. (2008) address that on a high-quality platform users are given timely and accurate information, as well as a consistent and easy-to-navigate interface, that enables an interactive and responsive experience during transactions.

> *'Because energy trading has not been widely adopted, clear instructions and easy-to-understand energy usage reports are the key for new users to adopt the energy trading product'*.

The service quality of an energy trading platform, according to the interviewees, is essential for users to adopt a certain product.

> *'Users' concerns about service since energy trading is a new concept. When their energy supply is influenced by system failure, they hope to have someone from the platform solve the issues as soon as possible'*.

In addition, for the service quality, feedback is most concerned factors. As M4 explained:

> *'Special offers, loyalty programmes and similar activities are offered through social media platforms, and these channels can also be used to gather consumer feedback and understand their needs and expectations to gauge consumer sentiment and identify issues related to media and customer service'*.

This finding aligns with the study by Cronin Jr and Taylor (1992), illustrating that consumer studies indicate that users' perceptions of quality are critical to their intention to re-use the service. Thus, the questionnaire includes questions about how participants evaluate the information quality, system quality, and service quality of an energy trading product.

ii. Economic benefits

Cost, cheapness, and savings are widely discussed in relation to the core themes. Previous studies discussed that the cost can impact decision making on smart meters, as customers may be less willing to support smart-meter adoption if they are aware that enabling technologies would cost them more money up front (Krishnamurti et al., 2012). In China, farmers are highly willing to use the energy trading because they are subsidised by the government for complementary farming and light (M4 and M2). For example, M2 mentioned:

> *'Farmers now feel good about it because they are subsidised and can save money. What you cannot use can be sold back to the grid and charged in both directions.*

However, it is controversial whether the benefits of using clean energy will cover the initial investment. F1 explains:

> '*The initial installation cost is an important challenge for the implication of P2P energy trading, and the customer will consider how many years it will take to recover their investment'.*

Thus, P2P energy trading can reduce the energy cost, but there is an initial installation cost which customers need to be aware of before implication. It was identified during the interviews that there are upfront investments which customers are not comfortable with, such as the cost of solar panel installation, as they are unsure of the time needed for it to pay for itself.

> *'Time differs in the UK and the Netherlands. In the Netherlands, several customers no longer require an upfront investment for participation in the energy trade. Customers in the UK are not as enthusiastic about using them, and although energy companies offer discounts and promotions, there are concerns about whether the cost of installation will pay for itself*'.

Secondly, there is the element of financial loss which can influence product selection but is a relatively new issue in comparison to previous literature. The main concerns about energy coins are how they will be used, whether they will be safe, and whether they will retain their value as well as traditional currencies. As M3 explained:

'*E- coins built on Ethereum are subject to price adjustments as the price of Ethereum goes up and down. However, the price of energy does not fluctuate as much as that of Ether, which creates a price differential. Much of this price difference is unacceptable to consumers.*

The volatility of e-coins is a serious challenge to the adoption of energy trading products, though cryptocurrencies may improve the security in transactions and data exchange (Bunjaku et al., 2017; Feng et al., 2019; Ivashchenko, 2016; Kowalski et al., 2021).

iii. Sustainability

As a motivated factor, environmental benefits are different for organisations and individuals in different culture backgrounds. M2 explains that in China it is mainly faced with the agriculture and the government can provide certain allowances. For the individual trader, it is quite difficult. While, in the Western countries, the trading object can be both, such as the organisations in the UK and individuals in the Netherlands. It is noteworthy that the government has been encouraging P2P energy trading with technical support. Various joint ventures exist which involve universities, energy enterprises, and governments. These are designed to attain flexibility in relation to energy supply whilst ensuring that environmental cleanliness is maintained.

In addition, organisations and individuals have different motivations for energy trading. For example, companies tend to be intrinsically propelled by the use of new energy sources. As F2 explained:

'*Many of our clients are not looking for cost savings, they are looking for a cultural* *fit for their company. For example, some companies have a green culture, so they prefer to use clean energy when choosing a power company'.*

Aside from price, a major reason for them to begin trading is energy conservation and the desire to ensure corporate value. From an utilitarian perspective, collaborative consumption creates sustainable means to consume and benefit many people rather than a few individuals (Binninger et al., 2015). The utilitarian theory shows collaborative consumption to be a form of CSR because it allows society to benefit from what they do not have.

iv. Stable supply

Customers in the UK lack the necessary amount of electricity generated from their roofs. This suggests that there is an unstable supply of electricity due to weather conditions. F2 stated 'T*his must be merged with the energy carried by the main grids*. Regarding the grid's setup, electricity supply in the UK is dominated by the state grid, the Northern grid, and the Southern grid. In China, the distribution of generated power is based on agricultural solar projects. Any surplus generated power is then merged into the main grid. Thus, F1 explained:

'Some customers will ask if there is a battery storage guarantee for this two-way deal, because when they do not have enough storage of their own, they won't sell anymore. They need to know exactly how much they have left. A stable storage and supply are what makes them dare to use it'.

In European countries, however, the main power grid remains separate. Each provider can make their own decisions concerning the flow direction regarding where the energy is generated and its destination.

5.4.2 Epistemic Value

Novelty seeking can impact the selection of energy products with the premise of lower cost. This is a new finding supplementary to previous literature. Customer acceptance of blockchain products depends on the customer's ability to innovate, but this innovation receives an influence from the price of energy. The personal innovativeness refers to a person's proclivity to accept new items, such as technologies, products, or services, before others (Lu et al., 2005; Rogers, 2002). For example, M3 mentioned:

> *‘Some customers want to try something new that they don't know about but, after giving them the knowledge, they continue to consult the price issue’*.

This finding aligns with the theoretical and empirical studies exploring the influence of consumers’ perceived epistemic value on their intention and behaviour (Assaker et al., 2011; Hirschman, 1980; Jang & Feng, 2007).

5.4.3 Social Value

Social influence and community belonging were found to motivate the P2P energy trading under the social motivations from the interview. Firstly, social influence can affect the selection of energy trading products via Bandwagon Effect (Schmitt-Beck, 2015). M3 explains that ‘*with the availability of financial assistance, your installation of this solar panel will cause your relatives and friends to want to install it as well’.*

In reflection of the literature review, social influence can be introduced through culture, nationalities, family, friends, and relatives (Moutinho, 1987; Sedera et al., 2017). If there is a neighbour or relative who used a P2P platform it might influence your decision on P2P energy trading.

Secondly, community belonging can motivate P2P energy trading, but community concern is also an important code in the set-up categories. What traders consider is whether their neighbours used the platform or whether the trading gives them a sense of belonging in the community. Both M1 and M3 support this argument, with M1 illustrating that ‘*Consumers are more concerned about popularity, if my neighbours have them, do I have to install them, is it better to be integrated in a community’*. As mentioned in the previous literature, it is believed that community membership or seeking to be included in a community or group is a factor that leads to the practice of sharing or CC activities (Galbreth et al., 2012; Neilson, 2010).

5.4.4 Emotional Value

The feeling of uncertainty can impact product selection. Perceived risk means the risk of losing money when using an e-service to achieve a desired result (Featherman & Pavlou, 2003). In terms of product selection, consumers are mainly concerned about financial and privacy risk.

> *'Customers do not understand how smart contracts protect their privacy and they are concerned that their privacy will be compromised. Because certain products require payment using energy, they believe that cryptocurrencies are insecure and may make their payment prices too volatile for them to use blockchain products'.*

> *'Cryptocurrencies pose financial risks due to excessive waveforms. Customers therefore do not want to choose some products that use cryptocurrencies for payment'.*

Previous studies illustrate that consumers identify uncertainty and risks, such as financial losses, when evaluating a product or service purchase which may cause them anxiety and discomfort. This perceived risk therefore further influences their next decision (Im et al., 2008; Martins et al., 2014; Pavlou, 2003).

5.4.5 Conditional Value

i. Political ideology

The results indicate that political ideology can influence how people perceive energy trading platforms, so that different countries with different ideology adopt different approaches. This process may be split into two steps. Firstly, the government-controlled economy decides energy planning is centralised and participation in energy planning is limited, which may slow the P2P energy trading progress:

> *'Currently, P2P energy trading is mostly applied on the agriculture and large organizations, as Chinese ideology is different with western. the government believes the order is important, thus, the individual trading market is difficult'.*

In contrast, in the west, there is a clearer understanding of energy planning and energy legislation due to the free market:

> *'In the UK market, consumers may have to wait a while to get in touch with their energy suppliers because energy is a biproduct of a different purchase such as an electric vehicle*'.

As Sovacool and Valentine (2010) explain, China and India have more centralised governments who decide the power planning than other decentralised governments. In other words, the limited power planning decided the development of P2P energy trading progress is slower than western countries.

Political ideology also determines the way in which energy settlements are made. For example, the energy trading products in China may not use cryptocurrencies which are commonly used in other trading grids (Chaudhary et al., 2019; N. Wang et al., 2019).

> *'Cryptocurrencies are not allowed in China because of the damage they do to the financial marke*t'.

Secondly, those different practices by governments with different ideologies may shape the expectations of the local consumers and further influence those consumers' product selection (Vroom, 1964).

> 'Consumers in China may tend to choose the government-backed trading products

to lower the uncertainty of the market'.

These findings align with previous literature (Crockett & Wallendorf, 2004; Duhachek et al., 2014; Jung et al., 2017; Madani et al., 2021) and justify the inclusion of the question of political ideology in the questionnaire bon the adoption of energy trading.

5.5 Summary

This chapter presents the findings of six qualitative interviews and compares the different motivations to P2P energy trading with previous literature. Most of the findings align with literature but three new insights have been identified. Firstly, interviewees mentioned the concern of financial loss because of the depreciation of cryptocurrencies which is unique comparing with other contexts. Secondly, the feeling of uncertainty caused users to worry about normal energy supply in daily life. Finally, political ideology plays a key role in the energy policy of a government which may further form the expectation of local consumers.

Chapter 6: Phase 1 - User-Generated Contents Analysis

This chapter is based on feedback from social media, encompassing the public channel (Twitter) as well as the customer channel (customer reviews of Vandebron). The topic modelling analysis on Twitter uncovers the most common topics discussed in relation to P2P energy trading due to the extensive development of online technology and its growing public acceptance. A similar analysis on Vandebron can also reflect customers' feedback and arguments regarding P2P energy trading.

6.1 User-Generated Contents (UGC)

User-Generated Content refers to the content created by users without commercial interest and published on publicly accessible websites or social media platforms (Daugherty et al., 2008; Smith et al., 2012). Generally, UGC data includes online text data and online photo data. These UGCs appear in social media platforms such as YouTube, Twitter, Myspace, and Facebook. Also, UGC is transforming the world in terms of entertainment, communication, and information, owing to its self-sustaining nature and ever-increasing audience size (Shao, 2009). The UGC has been applied in different research, such as customer service (Chatterjee, 2019; Mastrogiacomo et al., 2021); education evaluation (Halees, 2011; Ray et al., 2021); and tourism behaviour analysis (K. Zhang et al., 2019; Zhang et al., 2020).

Analysing UGC data are very important for understanding the user preference (Stone & Choi, 2013; Susarla et al., 2012; Timoshenko & Hauser, 2019). The reason that this study selects the UGC analysis as it is the alternative way of analysing the motivations of P2P energy trading and can provide deeper insights than the other quality analysis, such as interviews in the collaborative consumption motivation analysis. It is easier to create a connection with others who have similar interests, which means it is easier to find the common interests among different persons by analysing the text data (Bowman & Willis, 2003); it can understand the consumer perception about the products or a service (Chatterjee, 2019), such as P2P energy trading.

6.2 UGC Data Analysis: Topic modelling

Topic modelling refers to computer algorithms in which latent patterns concerning word frequency are detected via word distribution in a set of documents (Deerwester et al., 1990). It is an NLP-based application for text mining and it aims to explore underlying structures or themes in textual documents collections (Eickhoff & Neuss, 2017). Berry and Castellanos (2004) define “text mining” as a process of discovering previously unknown information by automatically analysing unstructured written resources. The extracted information can be further used to generate new facts or hypotheses (Gupta & Lehal, 2009)

UGC analysis is to extract factors influencing P2P energy trading from the social media data with the topic modelling. The selected region only includes UK and United states of America, as the handling the English at a time will make the analysis more accurate using topic modelling. This is an additional supplement to the interview and the literature review. Topic modelling is important for this research is because it helps to discover what users are talking about, uncover underlying thematic trends and track them over time, and determine the most relevant publications for a given topic (Nikolenko et al., 2017). Using topic modelling can detect the topics generated from the social media what users discussed about the P2P energy trading during a period. This also provides additional evidence for designing the survey.

Figure 6.1 demonstrates the process of dealing with UGC. After the data collection, this research used multiple techniques to pre-process the textual data. In specific, non-English words and words that have no semantic information, including URLs, stop-words and other meaningless words (Tirunillai & Tellis, 2014). The pre-processed texts were then used to generate corpus which is required by Latent Dirichlet Allocation. During the corpus preparation step, the Phrases module in Genism (Řehůřek & Sojka, 2011) was used to make bigrams. Bigrams are word pairs that are used in a document representation to improve text categorisation results, and are able to provide more information than unigrams (Bekkerman & Allan, 2003; Bhakkad et al., 2013; Wang & Manning, 2012). Lemmatisation is another technique used to normalise the texts.

Specifically, this research applied part-of-speech tagging of NLTK (Loper & Bird, 2002) to retain only words that are nouns (Martin & Johnson, 2015). Then, the corpus was generated by the Corpora module of Gensim (Řehůřek & Sojka, 2011) using the lemmatised words.

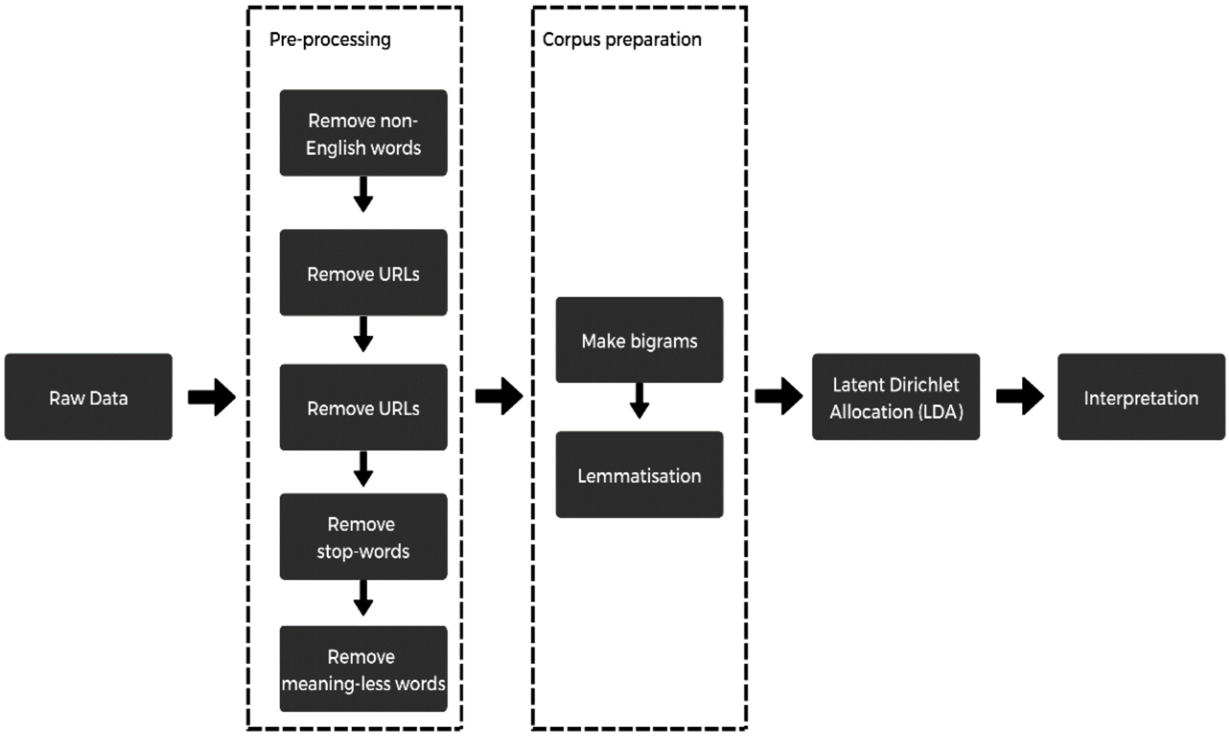


Figure 6.1 Data Process of UGC

6.2.1 Topic Modelling Process

Latent Dirichlet Allocation (LDA) is one of the most common and well-studied topic models (Blei et al., 2003). The output creates a set of topics related to word clusters that occur simultaneously in these documents based on specific configurations. Hence, to demonstrate these topics and interpret the P2P energy trading challenges, topic modelling was implemented to extract topics. LDA follows the Bayesian approach by assuming that the words included in the text are independently derived from a combination of baskets in which every basket includes a set of words in addition to the generative process concerning every tweet (Blei et al., 2003). In Figure 6.2, α denotes the per-document topic distributions, θ comprises the topic distribution related to document m, β indicates the per-topic word distribution, φ represents the word distribution for topic k, z signifies the topic for the n-th word in document m, and w stands for the specific word. The grey shading of W suggests that the only observable variables are words w_{ij}, whereas the other variables are latent. (Blei et al., 2003) recommend using a sparse Dirichlet prior to model the topic-word distribution, which is predicated upon the assumption that there is a skewed probability distribution related to words in a topic because there is high probability of there being only a small set of words. The acquired model is the LDA variant that is most implemented. The model's plate notation is illustrated on the right wherein K signifies the number of topics while φ_1,…φ_k stand for V-dimensional vectors that contain the Dirichlet-distributed topic-word distributions' parameters (V indicates the number of words in the vocabulary).

The variables indicated by θ and φ should be regarded as matrices derived by decomposing the original document-word matrix which denotes the corpus of the modelled documents. Therefore, θ includes document-defined rows and topic-defined columns, whereas φ encompasses topic-defined rows and word-defined columns. Hence, φ_1,…φ_k indicate a set of rows, or vectors, that are distributed across words, whereas θ_1,…θ_k indicate a set of rows that are distributed across topics.

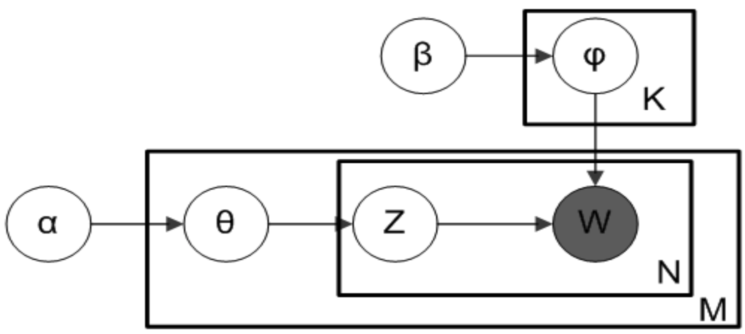


Figure 6.2. A plate diagram of an LDA model

i. Identifying Keywords

Google Trends provide new trends services that display the emerging issues based on their data. For example, Google displays the "top 10 most searched topics" every hour and those emerging issues reflect public interest trends (Han et al., 2012). In the current research, the researchers explored energy trading via Google Trends to identify keywords because Google Trends provides weekly information on keyword search frequency on the Google search engine. Using Google Trends permits a researcher to search volume patterns for keywords that can be analysed according to category and by the location of those making the search. Moreover, when we search the keywords "energy trading", relevant words such as "p2p energy trading" and "peer-to-peer energy trading" were additionally recommended. Hence, these three keywords also became the Twitter keyword searches.

ii. Determining Topic Numbers

After approximating the LDA posterior distribution, the K topics are represented as multinomial distributions over V. Each topic distribution contains every word but assigns a different probability to each of the words; the words within topics with high probability are words that tend to co-occur more frequently. These high-probability words, usually the top 10 or top 15, are used to interpret and semantically label the topics. However, many topics are defined by K, wherein a low K results in too few topics or very broad topics, whereas a high K results in uninterpretable topics or topics that ideally should have been merged. Choosing the correct value of K is, therefore, an important task in topic modelling algorithms (Syed & Spruit, 2017). Automated coherence score measures (Equation 6.1) were used for determining topic number K in this research, where $V^{(t)} = \left(v_1^{(t)}, \ldots, v_M^{(t)}\right)$ comprises a list of the M most probable words in topic t. A smoothing count of 1 is included to avoid adopting the logarithm of zero Topic Coherence measures score a single topic by measuring the degree of semantic similarity between high scoring words in the topic (Mimno et al., 2011).

$$C\left(t; V^{(t)}\right) = \sum_{m=2}^{M} \sum_{l=1}^{m-1} \log \frac{D\left(v_m^{(t)}, v_l^{(t)}\right) + 1}{D\left(v_l^{(t)}\right)} \quad (6.1)$$

6.2.2 Data Extraction Process

In this research, social media was selected as a channel for evaluating the factors influencing P2P energy trading because social media has drastically altered the ways in which people obtain and transform information since Web 2.0's launch. Oh et al. (2020) state that social media is a form of communication that allows information to be obtained quickly and effectively. Tweets and Vandebron reviews are selected to be the data source of the topic modelling. This research focused on two channels, one of which considered public opinions from Twitter, whilst the other examined feedback from real Vandebron customers. Vandebron is a long-established P2P energy trading company. And in contrast to other companies, Vandebron has significant more user reviews online.

Two different channels were chosen because of the different nature of the two sources. Tweets are more about public opinions and the topics generated from tweets may not represent the opinions of the real users. While online comments from Vandebron were selected which is the real reflection of the users. By selecting topics from these two different channels, it is possible to understand the topics from social networks comprehensively.

In order to collect the data, ScrapeStorm utilised a command prompt for searching through the desired tweets by specifying the keywords, the start date, end date, language, number of tweets, as well as the output format (Taspinar & Schuirmann, 2017). It was also ensured that only actual tweet texts were collected and variables not relevant to assess sentiments including likes, retweets, and replies were excluded. Certain tweets also included website URLs that do not help with assessing feelings. Therefore, all of the URL links featured in the gathered tweets were deleted (Pak & Paroubek, 2010). Finally, 4464 tweets were crawled of which 4132 tweets were finally selected following the pre-processing of the data. Also, Vandebron reviews present the actual P2P energy trading experience of customers. For this, 2018 reviews on Vandebron were obtained from the Klantenvertellen.nl website through ScrapeStorm along with reviewer profiles on 1st, March 2020. And then, this research use python 3 to do the topic analysis.

The number of topics intended to be detected, K, is the only parameter required. Furthermore, the coherence score was implemented for establishing the most suitable

K value. The two social media channels include diverse topics. In this section, the results of the LDA topic model analysis of the tweet corpora for P2P energy trading are presented.

6.3 Results

6.3.1 Determining Topic Number

The Coherence Value of the topics was calculated. When the number of topics is 2, the Coherence Value is 0.3562. When the topic number is 8, the Coherence Value is 0.3788. When the topic number is 38, the Coherence Value is 0.3758, and when the topic number is 44, the Coherence Value changes to 0.3586. Also, as Figure 6.3 (left) illustrates, the highest Coherence Value is achieved when the topic number is 8. Thus, 8 topics were selected to be interpreted. For the data from Vandebron, the same procedures on the 1609 comments were conducted and finally chose K equals to 8 as well. This is demonstrated in Figure 6.3 (right), when K is equal to 8, the coherence reaches highest.

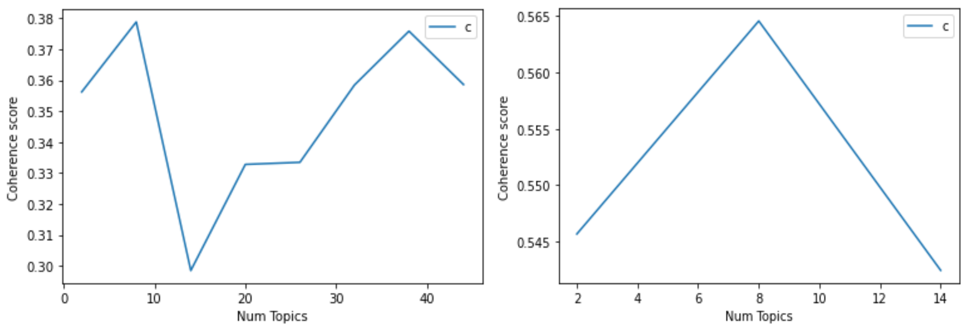


Figure 6. 3 Coherence value of topics of twitter(left) and Vandebron (right)

6.3.2 Topics Generated from UGC

It is worth noting that, even if the topic generation is automatic, the visualised results still require human power to interpret. Comparing with generating topics and subtopics, interpreting those topics can be the most challenging aspect of a study. Interpreting hierarchical topics is even more difficult compared with the analysis of general topics. The results for the generation of models, such as LDA, incorporate several dimensions including keywords and weight.

As the weight topics are demonstrated in the following Tables 6. 1, the top 15 most relevant terms under different topics were selected by adjusting relevance metrics (λ = 0.7). Twitter includes topics concerning the business mode such as the blockchain, payment, sustainability, economic benefit, system quality, service quality, and financial loss.

Table 6.1. Topics generated from twitter related with P2P energy trading

Topic number	Weight of tokens	Topic key words	Topics
1	16.3%	blockchain, power, platform, renew decentral, ledger, transact, technology, exchange, communion, smart, network	Blockchain
2	12.3%	wallet, digit, user, trust, kwhcoin, token, tokensale, spend	Payment
3	9.9%	solar energy, safe, green, interact, lend, transparent	Sustainability
4	9%	convert, connect, distribution, internet, passion, share, virtual, smart, storage, transfer, user, privacy	System Quality
5	6.7%	support, public, like, trust	Service Quality
6	5.4%	Price, sale, liker, lower, future use, increase	Economic Benefit
7	5.3%	digit, enron, currency, transfer,	Financial loss
8	4.8%	efficiency, solar, invest, wind, reliable	Sustainability

Topics extracted from Vandebron are different with the topics from Twitter. However, topic keywords from Twitter are more refined than topic keywords from Vandebron due to two reasons. Firstly, there are 4132 tweets on the Twitter topic modelling compare with 1609 customer comments from Vandebron. Secondly, tweets are broad opinions about P2P energy trading. While customer comments are more about the users' feedback after the trading. Thus, topics extracted from Vandebron as showed in Table 6.2 may have some degree of overlap and are more related to service quality. These topics are service quality, reputation, system quality, and emotion.

Table 6.2 Topics generated from Vandebron related to P2P energy trading

Topic Number	Weight of Tokens	Topic Key Words	Topics
1	28.8%	Zeker/definitely Bevelen/recommend Snel/quick Contract	Service Quality

		Prijs/price Prestatie/performance	
2	16.2%	Klantenservice/customer Goede/ good Prijs/price Betrouwbaarheid/ trustworthiness Verwacht/expect Eerlijk/fair Bereikbaar/reachable Vriendelijke/friendly	Service Quality
3	15.4%	Onaanvaardbaar/unacceptable Ster/star Leverancier/supplier	Reputation
4	9.3%	Communicatie/communication Altijd/always Betrouwbaar/reliable Geholpen/helped Energiebedrijf/energy company Tegen/against Misleidend/misleading Klantgericht/customer-oriented/switch overstap	System Quality
5	8.3%	Waardeloos/worthless Bezig/busy Helder/clear Antwoord/answer Snel/quick	Service Quality
6	7.6%	Zeker/definatetly Bevelen/recommended Overzichtelijk/clear Informatie/information Tevreden/satisfied Positief/positive Duurt/takes Lang/long	System Quality
7	7.3%	Redelijke/reasonable Prettig/pleasant Tevreden/satisfied Kwaliteit/quality Transparent contract	Service Quality
8	7%	Prettig/pleasant Persoonlijk/personal	Enjoyment

		Duurzame/sustainable Perfecte/perfect	

6.4 Discussion

Topics generated from twitter and Vandebron are different. Topics extracted from twitter are related with function values. Topic related with Vandebron are about the functional values as well as the social values and emotion value. The following section will explain these in detail.

6.4.1 Function Value

i. Net system benefit

Information quality, system quality, and service quality are the components of net system benefit which are the major topics related to P2P energy trading. Topic keywords under service quality can be negative, such as worthless, busy, or positive, such as quick, pleasant, and transparent. For example, some users note their positive interactions based on service feedback and information quality:

'The Vandebron employee was very friendly to me and explained everything very clearly"

"I have been a customer of 'Vandebron' almost since the very beginning and I have noticed the personal customer experience. I know where my electricity comes from. There's no big, bulky company between the supplier and the customer".

Thus, information quality is emphasised given its influence on decisions to use P2P energy trading. However, others note their disappointment at the service as evident through comments on the overall system quality:

'When I get the statement, I don't see the welcome bonus. And then I don't react anymore. Worthless service.'

'Via a vague link, I am stuck with a 3-year contract...ridiculous!' and 'it is ridiculous… I can't transfer my contract!'

These comments indicate the difficulties of system transmission and the system quality which impacts the use of P2P energy trading.

ii. Economic benefit

Economic benefit is the most discussed topic under the economical motivations on Twitter than Vandebron. For example,

"Our producers and consumers can directly agree on the price, which allows the consumers to lower their energy prices, and the producers to increase their earning potential."

"Co-ops in NC to offer efficient solutions to minimize energy cost for consumers."

Initial payment is a challenge even if energy companies claim their products can reduce energy costs. For example, *'reducing cost is good, what about the initial installation?'* This quotation confirms the impact of economic benefit in literature and the results from the interviews regarding the economic benefit (Bunjaku et al., 2017; Feng et al., 2019; Ivashchenko, 2016; Kowalski et al., 2021; Krishnamurti et al., 2012).

The benefit of using cryptocurrency is also discussed:

'Smart metres could quantify the amount of energy produced and consumed using solar energy, while energy trade and cryptocurrency payments would be governed by smart contracts and implemented on the blockchain.'.

'*I felt great about the SaitamaToken. it is a different kind of energy and style.*'

This indicates the benefits of digital currencies, and some members of the public are happy to test it despite the risk of currency depreciation.

iii. Sustainability

Green is an important topic under environmental motivations, which decide the intention of the P2P energy trading:

‘Opt out, his company is not really green, it’s an e energy trading company with little energy generation pretending to be green’.

Evidently, whether or not a product is green appeared to influence some users’ decisions to use the product. In addition, consumers tend to use green energy:

‘*I’m in U.K. but afaik* - *Europe has a large proportion of nuclear power and a network of interconnectors to allow energy trading and to balance the grid. Not sure about the states, I think they went gung-ho for green energy’.*

P2P energy trading is unique as one of its motivations is to increase trading efficiency and reduce environmental impacts. Therefore, the results of topical analysis align with previous literature on consumers’ expectations and their environmental concerns (ThØgersen, 1995; Thøgersen & Ölander, 2003).

6.4.2 Social Value

i. Reputation

Reputation is a popular discussed topic in P2P energy trading. However, the reputation was only extracted from the Vandebron feedback. *For example:*

‘A few weeks later, the news came that Vandebron had been taken over by Essent. Essent is a large energy giant with a poor ranking (see Greenpeace) and reputation. I now want to leave Vandebron and so I called at the beginning of December to cancel.’

Thus, the reputation of the platform is significantly important to continuous use of P2P energy trading. As previous studies note, in P2P platforms the functional factor of service levels and the symbolic factor of the service provider's reputation influence the customer behaviour process (Ert et al., 2016; Priporas et al., 2017). In other words, a solid reputation provides a basis of trust for the next exchange.

In addition, Puschmann and Alt (2016) emphasise that reputation complements ecological and economic benefits in the sharing economy, which contributes to consumers' social ambitions. As a capital asset, a positive reputation can take years to build yet can be destroyed almost overnight (McDonald, 2016). Hollowell et al. (2019) advise that customers become loyal to brands that they consider will have the most benefit to them in the sharing economy. The level of trust that people have about a product, person, company, or brand is influenced by both corporate and personal reputation, which enables people to work collaboratively.

6.4.3 Emotion Value

Enjoyment was apparent in feedback from Vandebron. For example, two participants claimed:

'*It is a pleasant experience in the communication with them, everything they do is about sustainability, or at least as sustainable as possible.'*

'They appear transparent to me. I am very happy with what they did.'

Consumers' acceptance of a new product or innovation is determined by their enjoyment (Ha & Stoel, 2009). Tussyadiah and Pesonen (2018) explain choosing Airbnb may be from consumers' sense of enjoyment.

6.5 Summary

Topics extracted from UGC are consistent with most previous studies. However, compared with the results from the interview, reputation was also extracted. A good reputation influences continuous trading intention. All of the topics that emerged in the UGC analysis are elaborated due to the topic keywords.

Furthermore, topics extracted from Twitter and Vandebron are slightly different. The topics for twitter lean more towards payment and blockchain because tweets are made by members of the public who are interested in this topic. However, the participants may not have experience in P2P energy trading, meaning tweets about P2P energy trading focus on technology popularity rather than actual experience.

In contrast, topics extracted from Vandebron are from real customers' feedback; the extracted topics are about the system quality, services quality, and the quality of information on the platform. Additionally, the reputation of the platform is extracted which reflects consumers' consideration when they choose a particular platform before trading. Thus, phase 2 integrates motivational factors from the interviews and UGC analysis with the literature review to conduct the survey.

Chapter 7: Phase 1 - Survey

Chapter 7 aims to build the dataset for solving the CCSR of ETRS. The dataset applied in the survey was collected internationally from China, America, Australia, the United Kingdom, Germany, and the Netherlands. This data provides different optimal solutions to Chinese customers and western customers on solving CCSR on P2P ETRS.

7.1 Survey Data Collection

The survey technique enables a large amount of primary quantitative data to be collected (Collis & Hussey, 2013); moreover, correlations between variables (two or more) may be examined, allowing hypotheses to be tested and findings to be generalised (Bryman, 2016). Surveys are used to provide a large amount of data in a short period of time, which in turn, provides further P2P energy trading data. Also, questions in the survey allow participants to express their attitudes and opinions (Hague, 2002). Crowther and Lancaster (2012) propose that research findings can also generate customer profiles, thereby improving the targeting of products and services and the evaluation of customer opinions, which enable researchers to assess the relations between products and users' preference on P2P energy trading to solve the CCSR on P2P energy trading.

In addition, this research collected two different datasets as different cultural predispositions to collectivism and individualism can impact the collaborative consumption that occurs during the exchange. To test whether there is a difference on the P2P ERTRS among different cultural backgrounds two international datasets are constructed. The datasets were named according to their geographical locations; one dataset is collected from China and the other datasets were collected from the United Kingdom, Germany, America, Australia, and the Netherlands; thus, this dataset was named the western dataset for the purposes of the study.

7.1.1 Sample Selection

This research considers the energy traders' preferences and attitudes to P2P energy trading; thus, the research participants are real customers with an experience of using P2P energy trading. As portrayed in Figure 7.1, P2P energy trading is discussed more in the United States of America, Canada, Australia, and European countries. In addition, a diverse population generally leads to precise statistical analysis because it enables more variables to be assessed simultaneously (Kumar, 2018).

However, Figure 7.1 also shows that China is not on the popular list because of China' access control to Google. In fact, P2P energy trading is already well developed in agriculture and fisheries in China, and it is now being piloted in some large cities between residential areas and commercial buildings, such as the recent pilot trading in Guangzhou (Ma & Shen, 2021). Thus, samples were collected from six different countries, including China.

Figure 7.1 The distribution of the interest of Peer-to-Peer energy trading topic over time (Google Trends, 2021)

7.1.2 Data Collection

The criteria indicated in the methodological literature were used to assure data quality in the investigation (Z. Lee et al., 2021; Steelman et al., 2014). The 'limitation of 1 attempt per HIT' function is enabled to identify duplicate responses, which flags an answer as a duplicate if respondents attempt to complete the survey using the same cookie. Moreover, responses that took less than five minutes to complete were eliminated. In addition, three attention check questions were designed throughout the questionnaire to catch any careless, random, or haphazard responses. The three statements required participants to select "yellow" for one statement, "strongly disagree" for another statement, "agree", and nothing else if they have been answering honestly so far’.

In addition, Web-based tools were utilised in this research including Amazon Mechanical Turk (M-Turk) and Wenjuanxing to distribute questionnaires and gather data. Web-based surveys represent an inexpensive, easy to conduct, and time-saving option. Nevertheless, the extent of the population is limited in a web-based survey because it invariably excludes those who are unable to access the Internet, which may lead to biased results (Collis & Hussey, 2013).

For the Chinese dataset, 980 responses were collected through Wenjuanxing, and 112 responses were removed due to being incomplete; thus, a sample of 868 responses was used for statistical analysis. For this platform, the age of participants was determined to be above 18 years who were given a reward of one pound upon answering all questions. The age group was established because of this age group’s relevance to the P2P energy trading in the household as well as considering the energy trading problem. For the western data collection, 800 responses were collected through M-Turk, and 208 incomplete responses were removed. As such, 592 responses were used for statistical analysis. Participants were asked to answer all questions within 20 minutes, after which they were given 1.75 pounds as a reward.

7.2 Questionnaire Design

Perceptual scales were used to measure the responses based on a 7-point Likert scale, as shown in table 7.1. The Likert scale is a prominent representative of summated rating scale that is widely used in survey research to give participants the ability to indicate whether they agree or disagree with a statement (Neuman, 2014). It provides an advantage in that it counters problems involving the development of pairs of dichotomous adjectives. Furthermore, it is treated as an effective method to measure and reflect participants’ attitudes toward an issue from different aspects and express opinions through one overall indicator (Kumar, 2018).

This survey comprises six components (as illustrated in the figure 4.1), namely functional value, conditional value, social value, emotional value, epistemic value, and selection of products.

To measure the functional value, system quality, service quality and information quality are adapted from the study (DeLone & McLean, 2004; Seddon, 1997) and stable supply is adapted from the study (Hahnel et al., 2020; Wang & Sun, 2017). Economic benefit is adapted from (Bock et al., 2005) , financial loss is adopted from (Abramova & Böhme, 2016), and sustainability is adapted from (Hamari et al., 2016).

To measure the social value, I used social benefit (Lin & Bhattacherjee, 2008), social influence (Bearden et al., 1989), community belonging (Möhlmann, 2015), subjective norms (Bock et al., 2005), and reputation (Doney & Cannon, 1997) as the key constructs. Moreover, participants' feeling of uncertainty (Kankanhalli et al., 2005) and enjoyment (Hamari et al., 2016), were the key concepts to measure the emotional value. Epistemic value was assessed using novelty seeking which is adapted from (Hirschman, 1980) study and variety seeking from research by (Kim & Jin, 2020).

Furthermore, questions related to demographic information were asked to assess the conditional value. These questions include gender, age, education, marriage status, family income, family size, and the number of children. Demographic information is a common approach in solving CCSR (Ávila et al., 2015; Safoury & Salah, 2013) and can impact perception to renewable energy trading (Ecker et al., 2017; Hahnel et al., 2020). In addition, questions related to building information were asked. These questions include the living area, bedroom numbers, house status, the age of the building, and the condition of the window glaze. I adopted the measurement items from previous studies such as (Belaïd, 2017; Rinaldi et al., 2018; Yu et al., 2011).

In addition, questions related to energy consumption information were asked. These questions include daily transport, 'Would you turn off the lights/computer/heating when going out?'; 'Would you wash the clothes at 40 degrees or less?'; 'Would you leave the mobile phone charger switched on at the socket when not in use?'; 'Would you boil the kettle with more water than you are going to use?' 'Who would you like to trade your energy with?'; and 'What kind of payment method would you like to use for energy trading?' These questions were adopted from measurement items in previous studies such as (Delzendeh et al., 2017; Paço & Lavrador, 2017; Zhao et al., 2019).

Cultural differences are measured by individualism/collectivism adopted from (Ianole-Călin et al., 2020) study. Political ideologies are measured by liberalism and conservatism which are adopted from (Caldwell et al., 2020) study.

Table 7.1: Measures

Construct	**Items**
INDIVILISM	I am a unique person/What others say cannot influence my decision /I can afford the decision made myself
COLLECTIVISM	I feel that participating in P2P energy trading soon is a personal obligation/I feel that I have a moral obligation to participate in P2P energy trading soon/I feel P2P energy trading will benefit the future generation and the world we live in.
LIBERALISM	I think the energy system changes are welcome/I think uncertainty in energy trading is acceptable/ I think fair distribution of energy is important
CONSTRUCTIVISM	I want the process of energy trading to be orderly and stable/I think it is necessary for the government to strengthen regulation of the transaction process/I think it is necessary for the government to strengthen regulation of the platform
INFORMATION QUALITY	I think the information provided by the energy trading website/app is reliable. /I think the information provided by the energy trading website/app is closed to the real situation. /I trust the information generated by P2P energy trading platform.
SERVICE QUALITY	I am aware of my welfare /When I confide my problem to my service provider, I know they will respond accordingly/I can count on my services provider and consider how their actions affect me
SYSTEM QUALITY	An easy to switch system is important to me /An easy to close account is important to me / I must be able to get the balance back
STABLE SUPPLY	Increased battery storage can encourage me to participate in P2P energy trading /Stable supply can encourage me to participate in P2P energy trading / Integrated within the main grids can help me during P2P energy trading
ECONOMICAL BENEFIT	It saves me money/Helps lower my electricity bill /Benefits me financially.
FINANCIAL LOSSES	I can accept losses due to inability to convert energy coins to conventional currencies, or not at a reasonable price /I can accept losses due to counterparties failing to meet contractual payments or settlement obligations /I can accept losses due to market volatility

SUSTAINABLITY	It helps to reduce the negative impact on the environment/Helps to reduce the consumption of energy and other resources/Allows me to be a more socially responsible civilian.
FEELING OF UNCERTAINTY	Using energy tokens increases my feeling of security /Using cryptocurrency will not worry me/Energy tokens based on bitcoin have no risk
ENJOYMENT	Buying energy from others sounds exciting/ The idea of purchasing neighbours' energy is enjoyable/P2P energy trading is fun
REPUTATION	This supplier has a reputation for being honest/This supplier is known to be concerned about customers/This supplier has a bad reputation in the market
NOVELTY SEEKING	How willing are you to seek information? / Do you search for the new and different trading objects? / How willing are you to try new types of trading?
VARIETY SEEKING	P2P energy trading allows me access different sources/ Wide ranges of suppliers are available if I choose to trade things with others
SOCIAL BENEFIT	Helps me connect with others/Allows me to get to know people from different neighbourhoods/Allows me to develop social relationships.
SOCIAL INFLUENCE	Will you use P2P energy trading when most people in your office/classroom use P2P energy trading/Will you use P2P energy trading when most people in your community use P2P energy/The more my friends around me start using P2P energy trading services, the more I am willing to use P2P energy trading services.
COMMUNITY BELONGING	The use of P2P energy trading allows me to be part of a group of like-minded people /The use of P2P energy trading allows me to be part of a group of people with similar interests
SUBJECTIVE NORMS	Most people who are important to me would think that using the P2P energy trading is a wise idea/Most people who are important to me would think I should use P2P energy trading/My family who are important to me would think that using P2P energy trading is a wise idea

Finally, products selection is included in the last section of the questionnaire, which is the rest component of the relation matrix of mitigating CCSR except users' trading preference. Thus, the type of products of the predicants finally selects or intention to use in the future trading was asked.

According to five components of energy trading, different products can be categorised in the market including hybrid platform, retailing platforms, community platforms, blockchain platforms, and vendor platforms (Morstyn, Farrell, et al., 2018; Sousa et al., 2019). In terms of the retail platforms, such as Piclo, lower prices provide a competitive edge. Electricity users, for instance, can make direct and cheap purchases in the wholesale market. Because of this, retail consumers gain the advantage of a competitive production market regarding commodity price and do not have to deal with promotion, customer service, and advertising expenses (Joskow, 2000).

Vendor platforms, such as Power Ledger, aim to help DER vendors improve their products' value. On the one hand, vendor systems reputation often causes significant challenges. For example, Alirhayim et al. (2017) suggest the reputation of products increases the likelihood of being selected. It is possible to reduce the functional loss related to high-voltage electricity by providing the specific battery system and PV panel vendors through battery storage and PV panels innovations (Guerrero et al., 2017; Morstyn, Farrell, et al., 2018).

The Community platforms such as Brooklyn microgrids, largely focus on identifying the benefits of separating microgrids from the main grid. At the community level, energy projects can focus on a common resource or goals such as improving supply–demand coordination for ensuring greater energy efficiency of renewable energy (Seyfang et al., 2014; Smith et al., 2016). DERs coordination is also considerably important for ensuring that supply is not hindered when separating microgrid from the main grid. Common goals such as improving pollution at a local level can encourage prosumers to organise community energy schemes (Walker & Devine-Wright, 2008).

The blockchain platform was designed based on the block chain technology such as SunContract. Studies have focused on different forms of blockchain implementation concerning energy trading, and there has been increased attention to subjects concerning security and privacy related to decentralised energy trading (Aitzhan et al., 2018; Li et al., 2018). Blockchain is an innovative technology involving transparency and ease of use that can effortlessly generate and consume energy (Alam, Li, & Patidar, 2015). Trading speed and efficiency can also be improved by using credit-based payment systems that will eliminate the need of blockchain using a reliable intermediary in P2P energy trading (Li et al., 2018).

The hybrid platform is the P2P market that has been integrated by community platform, the storage batteries, and national grids, such as SonnenCommunity. This type of platform provides greater system compatibility for a stable energy supply and affords greater predictability for grid operators (Sousa et al., 2019).

7.3 Descriptive result

7.3.1 Demographic Information Descriptive

From the Chinese dataset, as illustrated in Table 7.2, it can be found that most of the participants are in the 25 to 34 years age range. 81.8% of the participants have a bachelor's degree. 37.56% of the participants' annual income is between $20000 to $29999. 58.29 % of the participants are married and the most common family size is 1 to 3. While, from the Western dataset, 20.9 were from Australia, 10.1% were from Germany, 9.1% were from the Netherlands, 26.4% were from UK and 33.4% of the participants are from the USA. 52.87% of participants are male and 48% of them have a bachelor's degree. 18% of the participants' annual income is above $60,000. 59.8 % of the participants are married and the most common family size is 1 to 3.

The basic demographic information of the Chinese and Western user samples was found to be broadly similar based on the questionnaire information. Large salaries characterised most of the users, thus reflecting energy trading's character, which necessitates substantial preliminary funds. Following a level of education, energy trading was more tolerated by the audience.

Table 7.2. Demographic information profile of the respondents

	Frequency (Chinese/western)		Percentage (%) (Chinese/western)	
Age				
18-24	203	11	23.39	1.86
25-34	436	38	50.23	6.41
35-44	181	206	20.85	34.79
45-54	38	294	4.38	52.31
55-64	9	40	1.04	6.76
64+	1	3	0.12	0.51
Gender				
Female	416	268	47.93	45.27
Male	444	313	51.15	52.87
Prefer not to say	8	11	0.92	1.86
Education				
High school/collage	56	159	6.45	26.9
Bachelor’s Degree	710	284	81.8	48
Master’s Degree	93	135	10.71	22.8
Doctoral Degree	9	14	1.04	2.4
Annual Income (USD)				
Less than $20000	207	83	23.85	14
$200000-299999	326	97	37.56	16.4
$300000-399999	200	97	23.04	16.4
$400000-499999	76	96	8.76	16.2
$500000-599999	27	157	3.11	18.5
$600000+	32	109	3.69	18.4
Marriage Status				
Married	506	305	58.29	51.5
Divorced	6	19	0.69	3.2
Single	356	259	41.01	43.8
Family Size				
1-3	509	354	58.64	59.8
4-6	350	231	40.32	39
7+	9	7	1.04	1.2
Kids Number				
0	385	271	45.2	45.8
1	389	154	45.5	26
2-4	81	162	9.3	27.4
4+	0	5	0	0.84

7.3.2 Building Information Presentation

In the Chinese dataset, 91.36% of participants live in urban areas and 47.7% live in a three-bedroom house. 38.36% of the participants’ house age is 5 to 9 years and

53.69% of the windows are single glazed. While in the Western dataset, 79.6% of participants live in urban areas and 34.6% live in a two-bedroom house. 23.8% of the participants' house ages are 10 to 19 years and 79% of the windows are double glazed.

Concerning the different description, a marginal distinction was apparent, with the Chinese sample residing in dwellings built between five and nine years ago, whereas in the Western developed nations, houses built 10-19 years ago were resided in by 23.8% of the sample.

Moreover, compared with the Chinese sample the energy saving activities of Western users were far superior, for example with double-glazing being used by 79% of the latter. Regarding electricity consumption behaviour, self-driving was far higher among the Western database, thus contributing to the marked distinction between the Chinese and Western users.

Table 7.2 Building information presentation

	Frequency (Chinese/Western)		Percentage (%) (Chinese/Western)	
Living Area				
Urban	793	471	91.36	79.6
Rural	75	121	8.64	20.4
Bedroom				
1 bedroom	66	76	7.6	12.8
2 bedrooms	306	205	35.25	34.6
3 bedrooms	414	201	47.7	34
4 bedrooms	51	75	5.88	12.7
4+ bedrooms	31	35	3.58	5.9
Building Age				
Under 5 years	234	97	26.96	16.4
5 to 9 years	333	135	38.36	22.8
10 to 19 years	209	141	24.08	23.8
20 to 29 years	62	87	7.14	14.7
30 years above	30	132	3.46	22.3
Window glazing				
Double	466	468	46.31	79
Single	402	124	53.69	21
House Status				
Owned outright	386	208	44.4	35.1
Owned with a mortgage or loan	288	177	33.2	29.9
Rented from the council	8	121	0.92	20.4

Rented from a housing association	122	48	14	20.6
Rented from someone else	48	32	5.53	8.11
Other	16	6	1.84	1.01

7.3.3 Energy Consumption Information Presentation

In the Chinese dataset, 49.31% of the participants choose public transport. 96.66 of them remember to turn off the light when they go outside. However, the distribution of answers to the question 'boil the kettle with more water than you are going to use 'is relatively even. In the Western dataset, 60.1% of the participants choose self-driving and 76% respondents remember to turn off the light when going outside. However, 56% of participants said they "sometimes" 'boil the kettle with more water than is used'. 59.8% of the Chinese respondents would like to trade with the people they know, but only 38% of the western respondents would like to trade with family members.

Comparatively, public transport has a greater likelihood of being used by Chinese consumers, although greater cognisance of environmental sustainability is not necessarily the reason. Both databases had a comparable distribution of replies to the additional question concerning heating and switching lights off. Therefore, self-driving's difficulty due to heavy traffic and associated delays is the main reason for public transport use. A further reason is that compared with developed nations, per capita transport ownership remains limited because of China's nascent economic growth (Ma et al., 2019).

Table 7.3. Energy consumption information profile of the respondents

	Frequency (Chinese/Western)		Percentage (%) (Chinese/Western)	
Daily transport				
Public transport	428	117	49.31	19.7
Self-driving	388	356	44.7	60.1
Recycling	33	72	3.8	12.2
taxi	11	11	1.27	1.9
Coach/train	1	10	0.12	1.7
Sharing car	7	26	0.81	4.4
Would you turn off the lights when going out?				
Every time	839	450	96.66	76
Sometimes	26	140	3	23.6

Never	3	2	0.35	0.3
Would you shut down your computer when going out?	97		16.4	
Every time	735	335	84.68	56.6
Sometimes	116	219	13.36	37
Never	17	38	1.96	6.4
Would you leave the heating on when going out?				
Every time	164	155	18.89	26.2
Sometimes	127	265	14.63	44.8
Never	577	172	66.47	29.1
Would you wash the clothes at 40 degrees or less?				
Every time	82	193	9.45	32.6
Sometimes	201	347	23.16	58.6
Never	585	52	67.4	8.8
Would you leave the mobile phone charger switched on at the socket when not in use?				
Every time	398	219	45.85	37
Sometimes	293	242	33.76	40.9
Never	177	131	20.39	22.1
Would you boil the kettle with more water than you are going to use?				
Every time	213	115	24.54	19.4
Sometimes	301	334	34.68	56.4
Never	354	143	40.78	24.2
Trading objects				
People you know	519	225	59.8	38
People you don't know	51	294	5.88	49.7
I don't care	298	73	34.3	12.3
Payment method				
Local currency	811	472	93.4	79.7
Cryptocurrency	57	120	6.6	20.3

7.4 Summary

The main task of this chapter was to identify the preferred platform for an energy trader using multiple human behaviour factors through a questionnaire. In other words, the survey creates the original database for the P2P ETRS. Because of the existence of the CCSR in the current ETRS, the dataset must create the relations between new users and new items.

The survey described the results and differences on the building information and energy consumption between the two datasets. All the noted differences have a considerable impact on the final energy trading. For example, the different house ages and daily transport habits which decide the recommendation models in the following chapter may also be different. Thus, Chapter 8 examines different viewpoints and uses a mechanical learning approach to investigate different optimal trading models.

Chapter 8: Phase 2 - Tree-Based Energy Trading System

The aim of this chapter is to prove the effectiveness of the proposed user profile strategy by proposing and evaluating the recommendation model, reducing the amount of user input, and understanding a consumer's decision-making behaviour when they select their preferred product within TCV.

Thus, this chapter consists of two parts. First, the Decision tree based P2P ETRS is discussed, including data pre-processing, construction of the classification model, and system evaluation. Then, the ensemble model is used in the section to investigate the improvement of the performance for the classification. Thus, this chapter comprises two experiments. Experiment one comprised a decision tree, the intention being to determine consumer decision-making strategies during product selection and identify the factors traders must consider in relation to energy product preferences. Experiment two sought to enhance the classification of the P2P ETRS-based decision tree by examining the ensemble model, including traditional classification algorithms such as CatBoost.

8.1 User Profile Dataset Preparation

8.1.1 Data Pre-processing

Data pre-processing is employed for data mining, which is an acknowledged data discovery procedure (García et al., 2016). Whilst improved predictive power is correlated with additional data quantities, the raw data must experience the pro-processing phase. This phase enables data duplication issues to be managed and the data to be cleansed and filtered. It also facilitates format conversion and compression (Taleb et al., 2015). Pre-processing modifies data in accordance with data mining algorithm needs, thereby assisting in the processing of data which is otherwise unamenable to treatment. Data pre-processing incorporates multiple techniques, such as:

i. Data cleansing

Specifically, data cleansing comprises error identification, detection, and correction. According to Rahm and Do (2000), data cleansing also locates, corrects, deletes, and updates data which is incomplete, inaccurate, or irrational. Redman (1998) reported that, unless extreme preventative measures are taken, sizeable datasets frequently incorporate entry issues. Specifically, field entry errors are in the region of 5%. In the current study, all questions are compulsory, and no missing values were observed in the dataset.

ii. Outlier Detection

Chandola et al. (2009) state that the terms outlier and anomaly are frequently used interchangeably, however they are detected in different ways. In addition, different approaches are applied to high dimensional data and low dimensional data. This study employed isolation forest to explicitly isolated outliers, rather than evaluating normal points and regions through the allocation of scores to data points. Hence, the minority status of outliers and their distinct properties were harnessed (Liu et al., 2008).

iii. Normalisation

Z-score normalisation and max-min normalisation are typically employed for data normalisation to address the features of different scales. The researchers employed the Min-Max normalisation to change the data to specific values in a linear fashion. The minimum value for each feature is changed to 0, with the maximum value being altered to 1 with all other values being transformed into decimals between 0 and 1. Equation 8.1 indicates the min-max normalisation:

$$\text{Normalised}(f) = \frac{f_i - f_{\min}}{f_{\max} - f_{\min}} \quad (8.1)$$

iv. Sampling

Essential class distribution is potentially critical to effective classifier sourcing. The distribution frequently approximate equality for all classes, however this is not true for all applications. When the majority class markedly outnumber the remaining minority classes this created imbalanced classification, which is problematic (He & Garcia, 2009). Predictive accuracy evaluates classification algorithm identification of samples from majority classes.

According to López et al. (2013), there are three principal approaches: data resampling, algorithmic modification, and cost-sensitive learning. Resampling incorporates oversampling and under-sampling approaches to the management of dataset imbalance. Both over-sampling and under-sampling algorithms can equalise the samples. However, the resultant sampling data may not be valid for the tree based classifier generation (Peng et al., 2019). The principal issue with imbalanced datasets is that machine learning techniques tend to overlook the minority class, thereby impeding performance in a vital area.

SMOTE resolves dataset imbalance by oversampling the minority class, typically by duplicating examples. However, duplications fail to add new information. In preference, new examples can be artificially created from existing examples. Hence SMOTE is not useful to the dataset because the anticipated neighbours are less than the n sample. SMOTE is appliable where minority classes comprise minor disjunct and the training model demonstrates overfit (Fernández et al., 2018). Conversely, under-sampling tends to generate minority sample loss (López et al., 2013).

This study adapted the use of a hybrid sampling algorithm known as ALLKNN-SMOTE (ALL-SMO), which is founded on the ALLKNN technique wherein majority class samples are omitted and minority class samples are selected on a manual basis using novel, randomly generated samples between the two minority class lines. This expands the minority class samples rendering them proximate to the number of majority class samples.

8.1.2 Feature Selection

Reduced computation time can be reduced, augmented prediction performance, and enhanced comprehension of machine learning data or configuration recognition applications is achieved via feature selection techniques (Chandrashekar & Sahin, 2014; Guyon & Elisseeff, 2003). The current research selected relevant features to create classification models through the eradication of features that are either redundant or irrelevant, as when one of two correlated features can suitably illustrate the data.

The three major approaches to feature selection comprise the filter method, the embedded method, and the wrapper method. Specifically, the filter method involves different ranking methods using ordering to create variable selection. Appropriate ranking criteria score the variables, whereby a threshold eliminates variables which fail to demonstrate threshold correlation criteria and Mutual Information (MI). This ranking system furthers appreciation of feature pertinence. This limitation is overcome by developing a non-linear fit for targets with single variables, permitting ranking in accordance with goodness of fit. The overfitting risk means that the use of non-linear pre-processing, including squaring and taking the square root, can be alternated with a simple correlation coefficient (Guyon & Elisseeff, 2003). MI considers the dependency measure of two variables. According to Shannon (2001), MI evaluate one random variable's dependence on data pertaining to another variable. Moreover, MI helps gauge the similarities between defined independent and dependent variables. The MI value is zero when two variables are mutually independent. More significant dependent variables correlate with greater MI value.

The current research used a two-step MI-based method of filtering, applying a Max- Relevance feature selection algorithm (Peng et al., 2005). Hence, the researchers selected MI as the metric to eliminate extraneous features, calculating the MI between the dependent and independent variables. These were ranked in descending order, using a manually selected threshold value to eradicate features that added less or were irrelevant to predictive power.

The researchers adopted the Minimum Redundancy Maximum Relevance algorithm (mRMR) for the second stage to compute both redundancy for each feature pair and relevance between class and features as per Equation 8.2 (Peng et al., 2005):

$$MRMR = \max_{1 \in \Omega_s} \left[I(i,h) - \frac{1}{|S|} \sum_{j \in s} MI(i,j) \right] \tag{8.2}$$

8.1.3 Model Selection

Estimation performance is critical since it predicts unseen data performance. The prediction of future data is problematic because it requires machine learning applications or the creation of new algorithms. It is challenging to decide upon the most appropriate model for all datasets. Therefore, it is preferable to avoid seeking a nonexistent perfect model. Rather, it is more useful to opt for an insightful model that approximates reality. P2P ETRS as a new system has a lack of multimedia data. Under this circumstance, the researchers collected structured data (in the survey) in the form of shallow learning models rather than deep learning models. Besides, which features impact the P2P ETRS and how they operate needs to be considered. Hence, the most probable prediction label must be sought. As such, the question of which label correlates with unseen data is resolved. Most approaches do not explain why the model predicts a given label or which features are more influential, bar the tree algorithms (Baehrens et al., 2010). In addition, tree models are one of the easiest models to interpret (Hara & Hayashi, 2016; Molnar, 2020).

The first experiment has two objectives: to appreciate the significance of different features according to product choice, and to employ the tree-based model to recommend optimal energy product-choice models. The decision tree model offers superior interpretive qualities enabling diverse feature impacts on the recommendations.

The second experiment addresses how the recommendation accuracy rate can be improved using CatBoost. According to Polikar (2006), ensemble approaches frequently elevate predictive performance including overfit aversion. Given the limited data, the learning algorithm locates multiple diverse hypotheses that accurately predict the training data yet fail to make useful predictions regarding unseen occurrences. The likelihood of an incorrect hypothesis being chosen is reduced by averaging different hypotheses, thereby elevating general predictive power.

CatBoost was selected on the basis of its feature combinations and its unbiased boosting in relation to categorical features, since the latter were amenable to the combination as new features. CatBoost employ a greedy approach to new combinations during the creation of new tree splits. For the first split, no combination is considered. However, subsequent splits find CatBoost combining all present combinations using the categorical features in the dataset. All splits are regarded as categories with two values to be employed in combination. In terms of unbiased boosting with categorical features, the target statistics methods can transform categorical features into numerical values, causing the distribution to deviate from the initial distribution, evading to a deviation in the solution. This issue is unavoidable with conventional GBDT approaches. Hence, the researchers made comparisons with alternative algorithms, including the decision tree employed in experiment one and XGBoost, AdaBoost, and RF.

Nevertheless, there is no guarantee that the selection of the model founded a fit index value and the addition of penalty related to the number of parameters will constitute a reasonable model; identifying a reasonable model is only possible when parsimonious models are considered. Human judgement alone cannot achieve this and must be overruled by mechanical data analytic procedure that evaluates the model's fit (Browne, 2000).

8.1.4 Model Evaluation

8.1.4.1 K-fold Cross Validation (CV)

During the model selection process, the training set creates the model. Meanwhile, the classifier hyperparameters are altered to confirm model stability and avert model overfitting to the training data. By applying K-fold CV, it is possible to augment the accuracy of predictions and retain bias (Kim, 2009; Molinaro et al., 2005). This approach was used at both the model regularisation and evaluation stages. VC is otherwise known as rotation estimate and comprises an adjunct to the hold-out approach. Moreover, it augments training data. The simplest cross-validation technique commences with two folds, within which the dataset is separated into a training component and a testing component. Subsequently, the test dataset is exchanged for the training dataset.

The complete dataset for the K-fold cross-validation is divided equally into K partitions. To process the valuation, diverse K-1 partitions were employed as training data, the rest of the partitions being used as validation data. Equation 8.3 illustrates the error as follows:

$$E_k = \sum_{i \in k} (y_i - p(x_i)) \tag{8.3}$$

where $p()$ denotes the model prediction, k signifies the k^{th} part of evaluation data. Once the errors are taken into account, it is possible to develop the cross-validate error as per Equation 8.4:

$$CV = \frac{1}{K}\sum_{k=1}^{K} E_k \tag{8.4}$$

In each iteration, the dataset was evaluated via the section of one-fold/partition. The remaining partitions were used in the training dataset. This procedure was reiterated *k* times. The *k*-fold cross-validation with a 5-fold and 10-fold cross-validation is the most popularly applied. When the choice of fold numbers is made, a larger *k* produces reduced bias and increased model variance. Leave-one-out is the most comprehensive computation method. When employed with k-fold cross-validation, it increases enhancement through the establishment of k = N, wherein N denotes the number of samples within the dataset (Hastie et al., 2009). The precision of the model is evaluated as the accuracy average of the k model. It was deemed desirable to use 5-fold cross-validation for the hyperparameter optimisation in the current study. However, with smaller samples, the leave-one-out cross-validation would be suitable.

8.1.4.2 Classification Accuracy

A P2P energy trading environment must first be built or simulated that has public data gathered from the online survey. A new system framework was then recommended to resolve the existing energy trading systems' weakness as well as enhance the trading experience. For determining the model best suited to the new system, various criteria had to be applied for assessing the model performance such as the precision, multiclass classification accuracy, F score, ROC curve, and recall. Typically, a model that has the highest accuracy rate or the lowest error rate is preferred. Accuracy or error rates by themselves do not confirm that the test model will perform efficiently, and other measurements can also help compare the different models' performances. There are various measures, but no perfect measure has been identified that can be regarded as ideal in all situations (Stehman, 1997). It is necessary to construct or simulate a P2P energy trading environment in accordance with public data extracted from the online survey.

a. *Confusion matrix*

The confusion matrix comprises the raw output produced by the classification model; the output indicates correct and incorrect classification. Table 8.1 presents the confusion matrix for multi class classifiers, wherein the first row in column one (Class A) comprises the True Positive (TP) value for Class A. The confusion matrix diagonal values indicate the TPs for columns two, three, and four.

Table 8.1 Examples of confusion matrix

	Class A	Class B	Class C	Class D
Class A	$TP_{AA(1,1)}$	$e_{AB(1,2)}$	$e_{AC(1,3)}$	$e_{AD(1,4)}$
Class B	$e_{BA(2,1)}$	$TP_{BB(2,2)}$	$e_{BC(2,3)}$	$e_{BD(2,4)}$
Class C	$e_{CA(3,1)}$	$e_{CB(3,2)}$	$TP_{CC(3,3)}$	$e_{CD(3,4)}$
Class D	$e_{DA(4,1)}$	$e_{DB(4,2)}$	$e_{DC(4,3)}$	$TP_{DD(4,4)}$

The ratio of the sum of the main diagonal values to the total confusion matrix values denotes the multi class classifier accuracy. Where cm represents the confusion matrix, classification accuracy Acc is presented in the following Equation (8.5):

$$Acc_{cm} = \left(\frac{\sum_{i=1}^{N} C_{ii}}{\sum_{i=1}^{N} \sum_{j=1}^{N} C_{ij}} \right) \tag{8.5}$$

where N comprises the number of classes, i signifies the row index, and j indicates the column index for the confusion matrix cm. The classifier Error Rate (ER) comprises the complement of the Acc, which is 1- Acc.

b. *ROC Curve*

Different measurements of scalar values, such as accuracy, sensitivity, and specificity, denote the classification performance. Whilst straightforward in application, classifier comparisons are problematic in that they are sensitive to imbalances in data and overlook performance in some classes. Receiver operating characteristics (ROC), Precision-Recall curves, and other graphical evaluation methods interpret classification performance differently (Tharwat, 2020). The curve denotes a plot which indicates classifier performance wherein TP is plotted in opposition to FP at different thresholds. Hence, diagnostic rule or classification performance can be shown (Swets, 1988). ROC can evaluation skew sensitivity and provide geometric information regarding the composition of metrics and the skew sensitivity (Powers, 2020).

c. Precision and F-score

In instances where TP is true positive, then TN is true negative, FP is false positive, and FN is false negative FN. Nevertheless, precision is indicative of how many predicted positive cases are real positives, as per the following equation (Equation 8.6):

$$\text{Precision} = \frac{|TP|}{|FP| + |TP|} \tag{8.6}$$

The F-measure indicates the harmonic mean of TP rate (TPR) and possesses positive predictive value since it demonstrated a combination of precision and recall (Buckland & Gey, 1994). The F-score represents improved accuracy in that it accommodates class discrimination. The maximum value of F-score comprises 1 and the minimum value of F-score is 0. The F-Score formula is presented in Equation 8.7, as follows:

$$\text{Fscore} = 2 \times \left(\frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \right) \tag{8.7}$$

8.1.5 Model Interpretation

Feature importance reveals which features impact the P2P ETRS and how they interact to influence P2P ETRS. This allows the researchers not only to introduce greater confidence to their model prediction but also to identify problematic behaviour (Hooker et al., 2018). Post-modelling interpretability illuminates the impact of features on prediction results and reveal how sensitive the model is to certain features. The methods available in this respect include partial dependence plots (Friedman, 2001), individual conditional expectation plots (Goldstein et al., 2015), and SHAP values (Lundberg et al., 2018). Previous studies have incorporated different classification algorithms to determine feature importance, including Grömping (2007) use of the linear regression framework to distinguish between different variance decomposition-based indicators, including dispersion importance and level or theoretical importance, in order to quantify reported variance or response changes for each regressor. Moreover, as noted by Pampel (2000), Logistic Regression provides a robust evaluation of how groups of independent impact a binary outcome. This is achieved by quantifying the unique contribution made by each variable.

Using these approaches allows the correlation between each feature and the modelled prediction to be quantified or envisioned in the absence of information about the target variable's true value. The researchers employed the SHAP value to interpret the models in the current study in order to gauge feature impacts in accordance with Shapley (1953) coalition game theory, as refined with generalisations and visualisations by (Lundberg et al., 2018). Datta et al. (2016) also conducted similar research, not only postulating a generalised concept of interest quantity for the characteristic functions of the Shapley value, but also emphasising the joint and marginal influence of feature sets. Most algorithms, bar plot visualisations, fail to explain outcomes. However, it is preferable to acquire an appreciation of the input features the obliged to nonlinear machine to produce its results for each separate data point, as observed by (Baehrens et al., 2010).

8.2 Decision Tree-Based P2P ETRS

DT can generate a pattern that can effectively clarify the options and provide information about their potential outcomes. Moreover, as highlighted by Waheed et al. (2006), they are helpful in deriving an objective overview of the risks and benefits of each option.

This study firstly puts forward a DT based P2P ETRS, which entails a two-step filtering process for method of selection. This eradicates redundant and irrelevant aspects and utilizes a DT classifier for consumer decision-making that brings interpretability, transparency, and efficiency to the process. In order to support high scalability, a massive collection of real-world data from two diverse case study regions is employed for system evaluation. This research has two primary aims: firstly, to examine the various features in the P2P ETRS; and secondly, to construct superior decision choice models. The suggested machine learning techniques were used in this study to determine energy platform choice processes that are not yet understood. SHAP values were produced from the models to illustrate the output classes to facilitate interpretation for the decision-maker.

8.2.1 Representation of the Datasets

As mentioned in Chapter 7, there are six components including functional value, conditional value, social value, emotional value, epistemic value, and selection of products. Except the demographic information, building information and energy consumption information under the conditional value, all the other features' groups are categorical features. The 7 scale (ranking from 1 to 7) questions are used to mark the trading behaviour, which decides its numeric character. The target column is the choice of the trading platform which includes 5 types and the distribution of the 5 class are illustrated in Figure 8.1. In this figure, numbers 1 to 5 represent hybrid platforms, retailing platforms, vendor platforms, blockchain platforms, and community platforms.

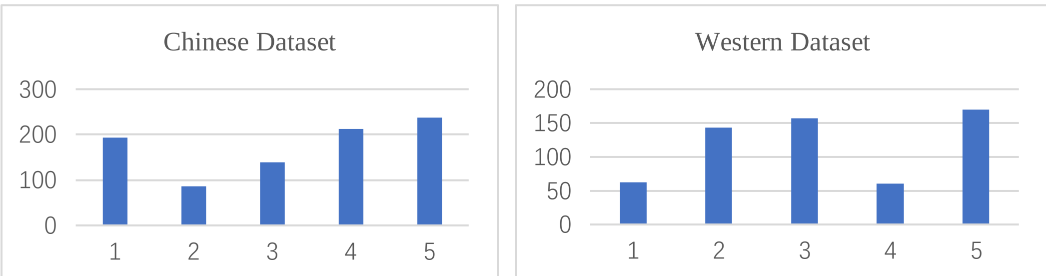


Figure 8.1 Distribution of the platform choice

8.2.2 Data Pre-Processing

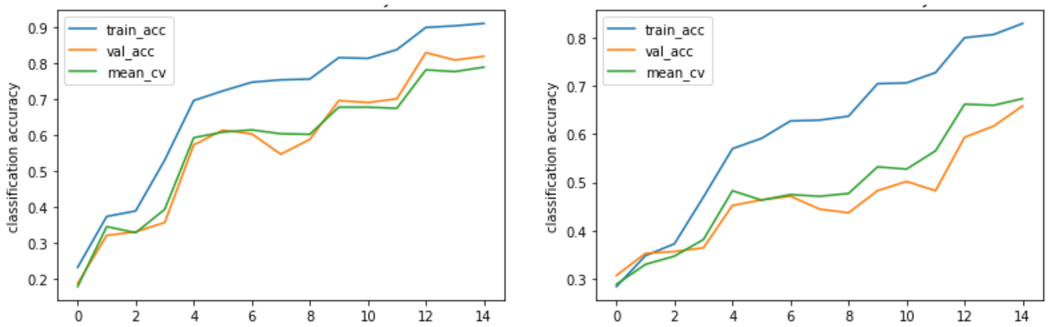
Following the application of the initial regroup, the min-max normalisation was employed to normalise the variables. Outliers and extremes are usually found in the data set; they need to be identified and removed to reduce the variance of the model. For the Chinese dataset, they were obtained during the data entry process by combining an automatic script to detect 123 outliers. For the Western dataset, on the other hand, 208 outliers were found by manual inspection. They were replaced manually by using the original values from the corresponding questionnaires.

Additionally, in the two datasets, the ALL-SMO approach was employed. Hence, equilibrium is established between the positive and negative class samples, as seen in Table 8.2. This technique tackles the issue of under-sampling, wherein excessive data is lost, and over-sampling problems are addressed, wherein excessive noisy data is produced. ALLKNN-SMOTE (ALL-SMO) both eliminates noise and duplication in majority class samples and results in balanced sampling through the addition of a limited number of effective minority class samples, thereby keeping valuable sample information pertaining to the majority class.

Table 8.2 Sample change from ALL-SMOTE

	Original imbalanced data	New balanced data
Number of minorities	128	385
Number of majorities	740	385
Percentage of minorities	14.74%	50%
Percentage of majorities	85.26%	50%
Total number	868	770

Also, as Figure 8.2 depicts, following the imbalanced process, the accuracy of the data classification is improved (Figure 8.2a) in comparison to its unprocessed state (Figure 8.2b). Hence, the resampling dataset with ALL-SMOTE is used.



(a) (b)

Figure 8.2 Accuracy on ALL-SMOTE imbalanced classification sampling

8.2.3 Feature Selection

To examine whether specific features aid in enhancing recommendation accuracy and reduce the user input in the user profile construction. Hence, feature selection was employed to assess if adjusting features can affect the classification accuracy. Following the cleaning and transforming of the dataset, the two-step MI-based filtering approach was applied (as detailed in 8.1.2), filtering out irrelevant features from the dataset. As Genuer et al. (2010) emphasise, the quantification of the variable significance is key for both ranking the variables prior to the application of a stepwise estimation model, and for interpreting the data and comprehending the underlying elements of many applied issues. Thus, in step one, a Max-Relevance feature selection algorithm was applied (as shown in algorithm 8.1) to filter out irrelevant features.

The MI score between each independent and dependent variable was calculated, and subsequently they were ranked in descending order and a threshold value of 0.1 was implemented to remove the features that made less significant contributions or were unconnected to predictive power. 0.1 is used as the threshold value rather than 0.2 because too many features cannot be removed at once. Figure 8.3 below shows that the same threshold was used on the two different datasets. The data sets removed features with an MI value lower than the threshold line, meaning that the remaining content of the two datasets were diverse.

Algorithm 8.1 Max-Relevance

Input: Discretized data d, class c

Output: feature set F

1: s = size(d);

2: **for** i = 1:s **do**

3: relevance(i) = mutual_info(d(:, i), c);

4: **end for**

5: **return** sort (relevance, 'descend');

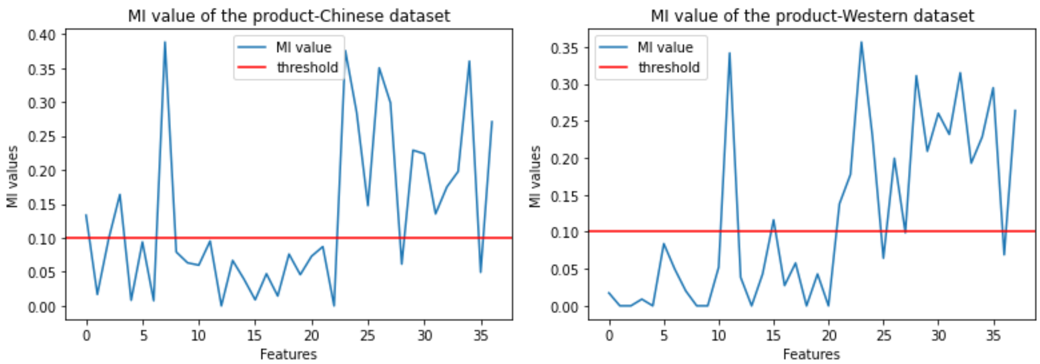


Figure 8.3 The MI values for each feature in each dataset

In accordance with Peng et al. (2005), in step 2 the minimum redundancy maximum relevance (mRMR) algorithm was applied based on minimal redundancy and maximal relevant criteria to filter out redundant features from the dataset. mRMR determines the redundancy for every pair of features and the relevance between features and class. In this study, MI was only considered for discrete variables and in mRMR form, as shown in algorithm 8.2.

Algorithm 8.2 Minimum Redundancy Maximum Relevance (mRMR)

Input: Discretized data d, class c, max number of features

Output: Selected feature set F.

1: s = size(d)

2: **for** i=1:s **do**

3: relevance(i) = MI(d(:, i), c);

4: **end for**

5: idx = sort(relevance, 'descend');

6: F(1) = idx(1);

```
idx_left = idx(2: max number of feature)
for p=2:s do
n = length(idx_left);
last_fea = length(F);
    for k=1:n do
        mi(p) = mutual_info(F, c)
        redun(idx_left(p), least_fea) = mutual_info(F, c);
        redun_mi(i) = sum(redun(idxleft(i), :)) / last_fea;
    end for
[G, F(p)] = max( mi(1: n) - redun_mi(1: n) );
g_mi(p) = G;
tmp_idx = F(p);
F(p) = idx_left(tmpidx);
idx_left(tmp_idx) = [];
end for
```

8.2.4 Classification and Model Construction with a DT

As the redundant and irrelevant features were filtered out, DT was selected as the classifier to build pertinent models. The justification for selecting DT models is that they facilitate superior interpretability; under the research objectives of this study, the goal is to comprehend how various features build the model. As discussed in the literature review, there are clear differences between decision tree algorithms. CART was deemed the most suitable classifier for this study, as one highly beneficial aspect of this algorithm is its cross-validation feature, as highlighted by Breiman et al. (1984), which can be used to identify overfitting. This is important, as not doing so is detrimental to future prediction.

Generally, there are two steps to build decision trees from examples. The first step involves training data being utilised to grow a decision tree, and the second step, as noted by Kim and Koehler (1995), entails pruning, which is the process of reducing the tree to prevent overfitting. Pruning is comprised of pre- and post-pruning processes. Pre-pruning involves methods of early stopping; for instance, the construction of the tree is stopped before it is fully grown. Mehta et al. (1995) describe post-pruning as growing the tree in full and subsequently trimming its nodes in a bottom-up manner. The pre-pruning approach is applied in this model based P2P ETRS. Cost-complexity pruning (CCP) was implemented by sequentially collapsing the node that generates the smallest per node rise in the error/cost, while simultaneously assessing the overall complexity (i.e., size) of the tree to ascertain the optimum pruned tree that minimises the cost-complexity function. CCP produces a series of trees T0…Tm, where T0 is the initial tree and Tm is the root itself. At step i, the tree is produced by taking a subtree from tree i-1 and inserting a leaf node in its place with value selected as per the tree construction algorithm. The pruning progresses by completing the creation of an ordered sequence and establishing the leaf node classes. When a tree T is pruned at node t, there is an apparent rise in the error rate $R(t) - R(T_t)$, and a reduction in the number of leaves $|N_{T_i}| - 1$.

Subsequently, as illustrated below in equation 8.8, the apparent error rate of each pruned leaf is calculated to increase. Conversely, the error estimate allocated to a subtree is the re-substitution error plus a factor a times the subtree size. The distinct a value that alter the tree size can be calculated using an efficient search algorithm. Additionally, the parameter is selected for the purpose of minimising the error on a holdout sample or utilising cross-validation. Following the identification of the optimal a value, the whole training set is used to grow the tree and it is then pruned with this value. In this study, the holdout method is used in the experiments, holding back 20% of the training set to estimate the optimum a parameter.

$$\alpha = \frac{R(t) - R(T_t)}{|N_{T_i}| - 1} \tag{8.8}$$

8.2.5 Dataset Setting

In this study's experiment, the Chinese and Western datasets are tested. These datasets encompass four diverse feature groups and 41 different features. To facilitate comparison, when the datasets were collected, the features were set in each one. Table 8.3 below gives a general overview of the features in question.

Table 8.3 Feature allocation for training

Feature group	Functional Value	Conditional Value	Social Value	Emotion Value	Epistemic Value
Number of Features	7	25	5	2	2

Scikit-learn 0.24.2 was used in this experiment to store and process recommendation models. The experiment was run ten times for each platform selection category, with an identical setup for each platform choice category. Randomised permutation and a stratified sampling method were applied to the dataset for each repetition to ensure homogeneity and heterogeneity within and between the strata, respectively.

Furthermore, a holdout sampling method was employed to divide the data into two partitions: 70% to be used as a training set and 30% to be used to test the models' actual performance. To ascertain the optimal parameters and evaluate the performance of the model, a stratified ten-fold cross-validation approached was used for the training set, which involved 5-fold CV. The minimum number of instances per leaf was set at 2. The classification accuracies of the training and validating sets of the various repetitions were averaged. The model with the highest mean validation classification accuracy and the smallest tree size, and that was not overtrained (meaning that the mean accuracy of the validation set was less than or equal to the mean accuracy of the training set) was deemed to be the optimum model.

Equation 8.9 below defines the decision trees' loss function. The decision tree grows recursively, and the recursive process finishes when it achieves maximum depth or when the purity of the data can no longer be decreased by splitting nodes.

$$J(k, t_k) = \frac{m_{\text{left}}}{m} G_{\text{left}} + \frac{m_{\text{right}}}{m} G_{\text{right}}$$

$$\text{where } \begin{cases} G_{\text{left/right}} & \text{measures the impurity of the left/right subset,} \\ m_{\text{left/right}} & \text{is the number of instances in the left/right subset.} \end{cases} \tag{8.9}$$

However, initially, for the hyperparameter searching, the pre pruned maximum depth is 3 to best avoid overfitting. Figure 8.4 depicts the change in decision tree platform with the maximum depth. When the maximum depth is 5, the testing score (represented by the blue line) rises with the increase in training points. It grows to a relatively high score of 0.85, thereby demonstrating the generalisability of the model. The training score (represented by the red line) reaches and remains constant at almost 0.92; thus, it is a good fit for the model and reaches a relatively high score. The testing score has two important stages where the rates of change differ.

Firstly, is the positive rate of change which endures until the maximum death is approximately 3, and secondly the rate of increase of the training points is very high until a maximum depth of five. Another is a maximum depth of between 5 to 10, where the rate remains constant, indicating that further training is not meaningful. Therefore, if the maximum depth is below 5, adding more training depth will certainly enhance the score, whereas if it exceeds 5, the rate plateaus, thus it is not particularly beneficial. In the Western dataset, the maximum depth is 4, and if it exceeds this, overfitting occurs. However, to facilitate ease of interpretation, the maximum depth was selected at 3, as when it is at 5 the tree becomes too big and complex for the decision-maker to comprehend.

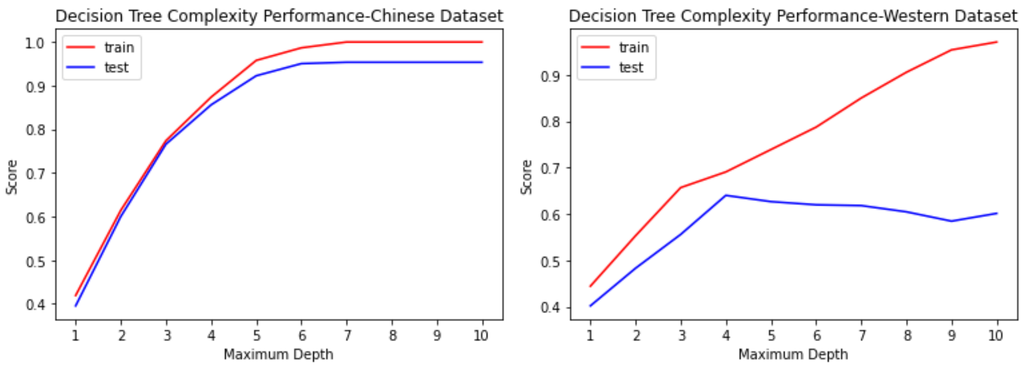
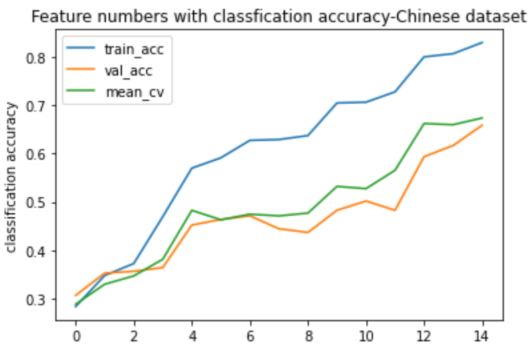


Figure 8.4 Decision Tree complexity performance

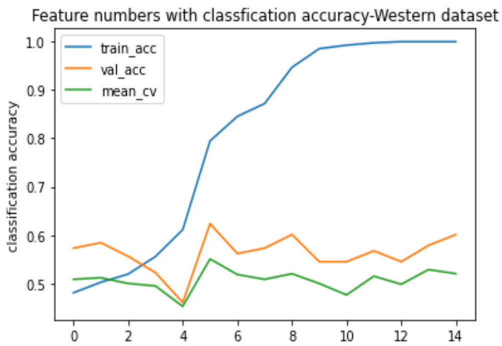
8.2.6 Experiment Result and Evaluation

Figure 8.5 below illustrates the classification accuracy of the initial n-selected features of each dataset. The classification accuracy of each dataset encompasses the mean classification accuracy of the training set, the validating set, and the test set with the minimised apparent error. In the Chinese dataset, it was found that the classification accuracy rate is significantly enhanced by amalgamating additional features.

In contrast to the Chinese dataset, when more features were added in the Western dataset, the accuracy rate remains between 50% and 60%. This is the case even when features continue to be added to the model, and these features do not offer any material relevance to the predicted class and the model is vulnerable to overfitting. Hence, the number of features selected is 5, while the training accuracy and test accuracy have the shortest distance.



(a)



(b)

Figure 8.5 Feature numbers with classification accuracy

In terms of the classification accuracy rate, the models' performance is assessed with the use of a confusion matrix, precision, recall, and F-measure. In addition, for the purpose of enhanced visualisation and interpretation of the models' performance, this study employs ROC curves and calculates AUC values. This is illustrated in Tables 8.4 and 8.5, which are the confusion matrices of both datasets. The bold font signifies correctly classified occurrences.

As shown in Table 8.6, it was found that the precision, recall, and F-measure is higher in the Chinese dataset than in the Western dataset. Furthermore, as Table 8.4 demonstrates, platform four was seriously misclassified in the confusion matrix, which resulted in the low classification accuracy. There are two primary reasons for this: firstly, the sample size of the Western dataset is inadequate, and secondly, the categories have a high ratio in the imbalanced class. Moreover, these models do not contain any associated significant features to classify the platform similarity.

Table 8.4 Confusion matrix for the test dataset - Chinese dataset

	Platform 1	Platform 2	Platform 3	Platform 4	Platform 5
Platform 1	50	1	0	1	2
Platform 2	0	22	0	0	1
Platform 3	2	0	43	4	0
Platform 4	1	0	0	49	6
Platform 5	1	1	0	21	56

Table 8.5 Confusion matrix of the decision tree - Western dataset

	Platform 1	Platform 2	Platform 3	Platform 4	Platform 5
Platform 1	10	1	9	0	0
Platform 2	2	32	1	1	4
Platform 3	2	2	43	2	6
Platform 4	0	5	0	10	2
Platform 5	0	2	4	3	37

Table 8.6 The precision, recall, and F-measure of each dataset

Dataset	Precision	Recall	F-score
Chinese Dataset	0.861040	0.842912	0.845494
Western Dataset	0.644916	0.685393	0.660597

For each dataset, the plot of the ROC curve shows the TP rate (sensitivity) against the FP rate (specificity). In addition, the plots show the area under the curve (AUC). In Figure 8.6, the classifier is not capable of discerning platform 3 from other destinations. Platform 3 shows an AUC of 0.48, which is an improvement of 0.28 on the random guessing. Platform 1 shows the highest AUC value in both the Chinese and Western datasets, at 0.94 and 0.86, respectively.

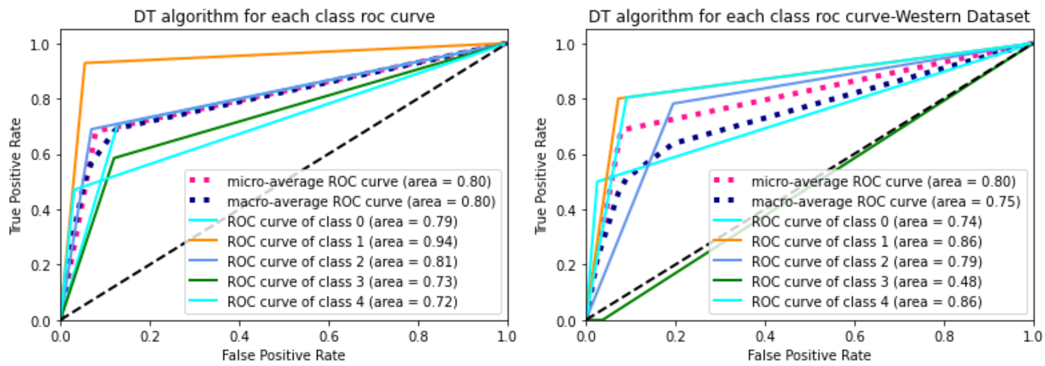


Figure 8.6 ROC curve summarizes the performance of the DT Algorithm

8.2.7 Model Interpretation

The SHAP value was used to investigate different features that contribute to the target column. Each model is explained utilising SHAP values to ascertain feature significance and the relevance of both the Chinese and Western datasets. SHAP values are calculated for each classifier using the KernelSHAP method (detailed in 8.1.5), which is available in the Lundburg-published SHAP Python package[1]. The same models were applied in a multiclass classification framework to enable the evaluation of the contribution of features for every class. In this case, each model is trained to predict one of the five platform classes.

From the global interpretability, the SHAP value can identify the contribution of each predictor. The vendor platform is represented by class 0, the retailing platform by class 1, the hybrid platform by class 2, the community platform by class 3, and the blockchain platform by class 4.

Figure 8.7 demonstrates that the feeling of uncertainty was most impactful to the vendor platforms, whilst also being significant to the community and blockchain platforms. House Status is shown to affect the Hybrid, community, and blockchain platforms. In addition, the vendor and retailing platforms are impacted by income.

[1] https://github.com/slundberg/shap

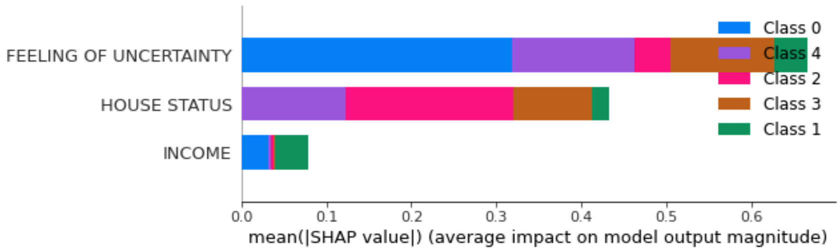


Figure 8.7 SHAP values plot explaining the DT classifier prediction of all classes - Chinese Dataset

Likewise, for the Western dataset, as shown in Figure 8.8 below, sustainability made the greatest contribution to the blockchain platform, whereas it seldom affected the community platforms. Additionally, income was significant to the retailing platform and vender platform. Novelty seeking plays a role in the blockchain and hybrid platforms.

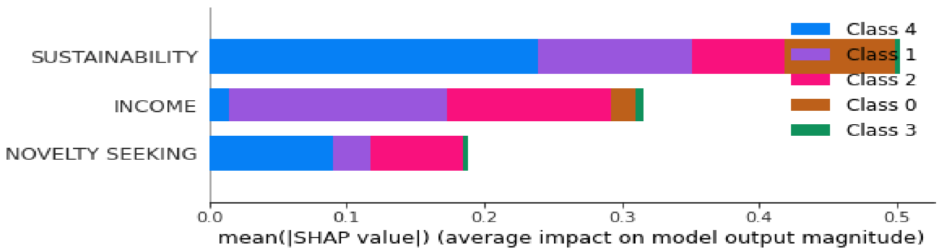


Figure 8.8 SHAP values plot explaining the DT classifier prediction of all classes - Western Dataset

8.3 Ensemble-Based P2P ETRS

8.3.1 Model Constructions

The second experiment focused on ensemble learning. The aim is to address whether ensemble learning can enhance the effectiveness of the classification performance of the DT-based P2P ETRS. The ability of an ensemble to generalise is usually substantially higher than that of base learners (Zhou, 2009). Ensemble learning, on the one hand, bagging can decompose the bias-variance, such as RF, which can considerably lower the variance, making it better suited to learners with high variance. On the other hand, boosting may considerably reduce bias while also reducing variance, therefore it is usually more effective on weak learners like decision trees. However, it is necessary to be noticed that while ensemble learning can potentially improve prediction performance and robustness, it is not guaranteed (Dietterich, 2000). Also, as discussed, overfitting is the most prevalent challenge that arises in the DT; therefore, in line with the tree model, a random forest (RF) algorithm and three gradient boosting variants, namely XGBoost, AdaBoost and CatBoost, were used to establish whether the model is optimal in this study.

An RF is comprised of numerous DT classifiers and utilises Breiman (2001) "bagging" concept to group many DTs into a single robust model. It employs the self-help method, which is referred to as bootstrap resampling technology, to create new training sample sets from the original training samples of N by selecting random k ($k < N$) sets of samples repeatedly. In comparison with other classifiers, RF utilises an amalgamation of classifiers which greatly decreases the rate of generalisation error and enhances the prediction capability of the model. It is possible that some samples may be collected more than once throughout the overall sampling process. In each round of the random sampling of bagging, approximately 36.8% of the training data will not be sampled, which Breiman (1996) calls out-of-bag (OOB) data. While this OOB data does not contribute to the model fitting during training, it can later be used to ascertain the models' generalisation capability.

In this experiment, OOB was used to calculate the models' generalisation. If the OOB error rate rises with the tree trained on both datasets, both trees will be removed, and new subsamples of the partially labelled dataset will be generated at random to train new trees. This method enables a subset comprised of too much noisy data to be disregarded and a new one produced; essentially, this means that noisy data can be identified and ignored. For this algorithm, two hyperparameters are defined: firstly, the maximum training times, and secondly, the bootstrap aggregating ratio. In this experiment, k-fold evaluation is used to ascertain the optimum values for each hyperparameter, and in most scenarios, the aggregation ratio was set at 0.6 and the maximum training times at 50.

In the RF, each feature chosen is generated at random from all m features, which itself has decreased the risk and propensity of overfitting. The model will not be established by eigenvalues or combinations of features. The higher level of randomness will not indefinitely enhance the control of the fitting ability of the model. Furthermore, in contrast to ordinary DTs, RF improves their foundation. According to G. Huang et al. (2019), with an ordinary DT, it is necessary to select an optimal feature from all m sample features on the node for the left and right subtree divisions of the DT. However, in an RF, each tree is a component of selected features. The best feature is chosen from these several features to separate the left and right subtrees of the DT, thereby increasing the randomness impact and generalisation capability of the model. Assuming that for each tree, m sub-features are selected, the smaller the m sub, the poorer the fitting degree of the model to the training set. While this increases the bias, it means that the generalisation capability will be stronger, and the discrepancy in the model will be reduced.

The general idea of the boosting method is to build a strong learner from a set of weak learners. Boosting functions by training a set of learners consecutively, and then amalgamating them for prediction purposes. Later, learners become stronger and concentrate on the errors of earlier learners to a greater degree. During the training stage, the initial weight of each training sample is equally allocated. For each round of boosting, the model is trained by the training set and error calculation is performed.

Next, the alpha value is used to update the weight. This process goes on until the last classifier has been trained. The weighted sum of the M classifiers is used to calculate the final model. Furthermore, the weight of the inaccurately classified sample rises. In this study, CatBoost, which supports multiclass problems by selecting the class that has the highest total vote, was applied. As per Ferov and Modrý (2016), base predictors are oblivious DTs in CatBoost. In this context, 'oblivious' refers to an identical splitting criterion being employed across a whole level of the tree. In this case, the trees are balanced, have a lower propensity for overfitting, and enable a much faster performance at testing time. Constructing a tree in CatBoost is similar to a standard GBDT process; however, Prokhorenkova et al. (2017) assert that categorical features maintains s supporting models M_r corresponding to TS based on σ1 , . . . , σs. Therefore, the last function of the CatBoost algorithm is expressed as follows:

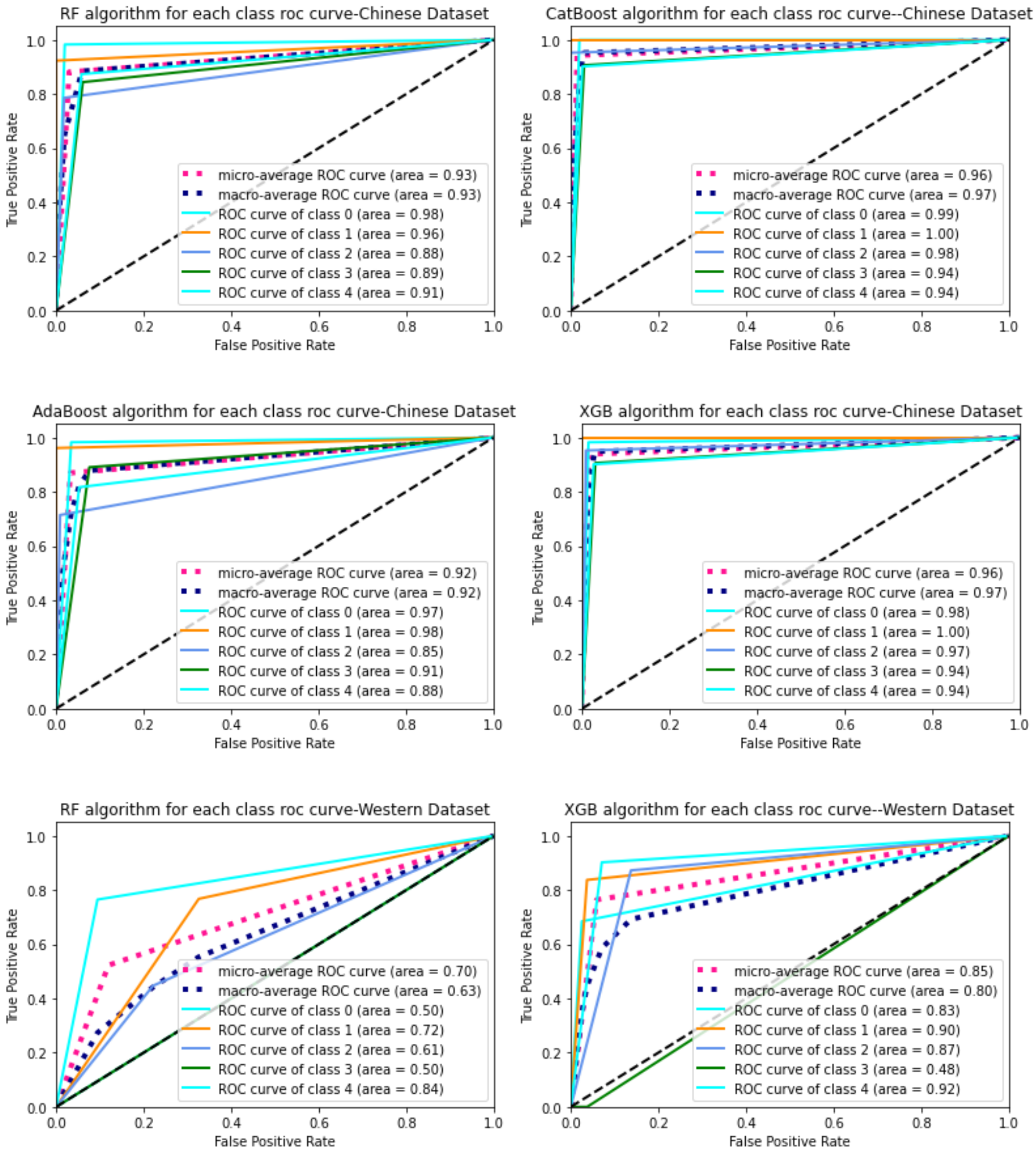
$$L_{\text{CB}} = -\frac{1}{M}\sum_{k=1}^{K}\sum_{m=1}^{M}\Big[w_{mcfn}^{k} \times y_m^k \times \log(x_{m,k}) + \sum_{k'=1}^{K} w_{mcfp}^{k,k'} \times y_m^k \times \log\big(1-(x_m,k')\big)\Big] \quad \text{s.t.} \quad k' \neq k \tag{8.10}$$

Where M is the number of training examples, K is the number of classes; y_m^k is the target label for training example m for class k; x_m is the input for training example m; w_{mcfn}^k is the marginal cost of a false negative over a true positive; and $w_{mcfp}^{k,k'}$ is the marginal cost of a false positive of class k' over a true negative, when the true positive is k. As they signify true negatives rather than false positives, the false positive matrix w_{mcfp} has vacant values along the main diagonal. In the context of probabilistic false positives, this matrix and its related triple summation in the loss function is the core of modelling additional loss.

8.3.2 Experiment Result and Evaluation

As depicted in Figure 8.9, the ROC curve plots show several aspects. Firstly, the classifier can discern all the class classifications in the Chinese dataset. Secondly, it cannot discern class 3 from other classes. Class 3 presents a lower accuracy of 48% in CatBoost, AdaBoost, and XGBoost. Class 1 shows the highest level of accuracy across all datasets. In this study, accuracy is prioritised over recall in the Chinese dataset, and vice versa in the Western dataset. Hence, classes that achieve high true positive rates while maintaining low false-positive rates, such as the destinations, are considered. Thirdly, there are comparable performances between XGBoost and CatBoost, which determined that the links between the number of trees and changes in the classification accuracy should be examined.

In this experiment, it is assumed that the numbers of trees in the RF are 20, 40, 60, 80, 100, 120, and 140. Hou et al. (2019) assert that typically, when the number of DTs in an RF increase, so too does the classification accuracy of the RF algorithm. The Chinese dataset results for this experiment are presented in Figure 8.10. Per the experimental results, with the exception of CatBoost, there is minimal fluctuation in the classification accuracy of the ensemble-based algorithm for both datasets as the number of DTs changes. When the number of DTs is less than 10, the classification accuracy of the CB and RF algorithms put forward in this study is lower for both datasets. For the Chinese dataset, the classification accuracy of the CB and XGBoost algorithms rises as the number of trees increases, and when the number of trees is 130, the classification accuracy is at its maximum. For the Western dataset, the CB accuracy is highest when the number of trees is 110, and the CB algorithms' classification accuracy is superior to the other three algorithms.



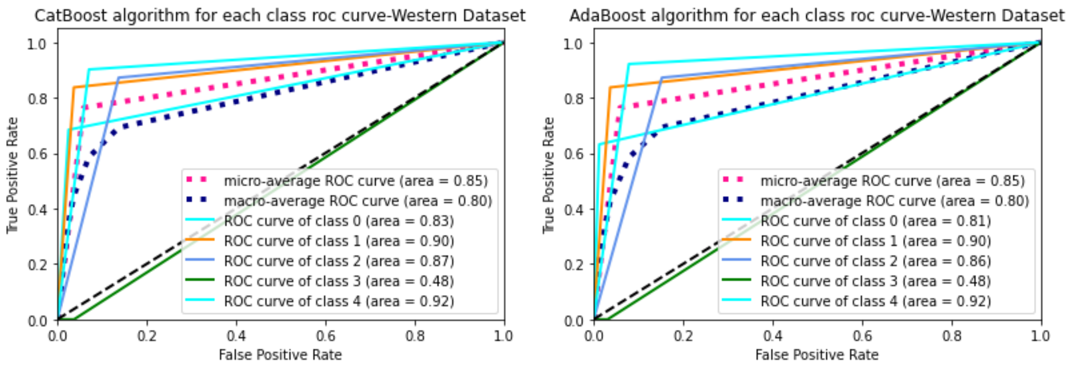


Figure 8.9 ROC curve summarizes the performances of the ensemble-based models

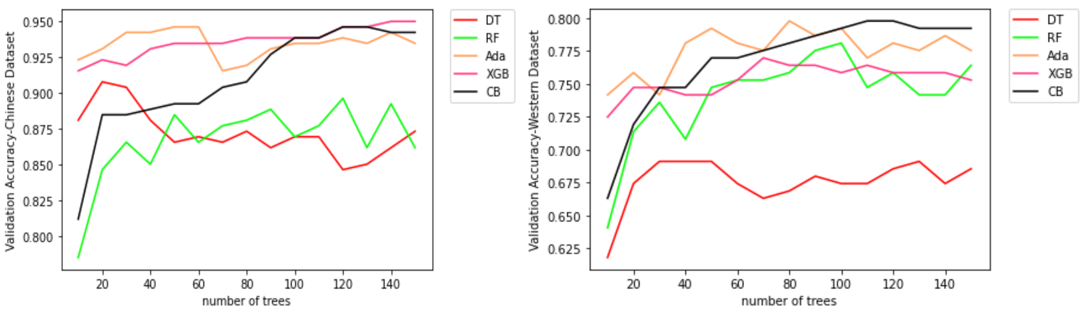


Figure 8.10 Impact of tree number on classification accuracy of the Chinese Dataset

CatBoost produced the best classification impact for the multiclass classification of all the ensembled models; therefore, CatBoost was selected as the best model. Loss errors were also tested with the increasing epochs. As depicted in Figure 8.11, when the number of iterations reaches 400, the model stops converging. For the two data sets, the loss function value decreases as the iterations increase. Evidently, the Western curves fluctuate more than the Chinese curves. As presented in Table 8.7, the accuracy of the Chinese dataset is 0.943 at 400 iterations, the depth is 8, and the learning rate is 0.1. In the Western dataset, the accuracy is 0.798 with CatBoost.

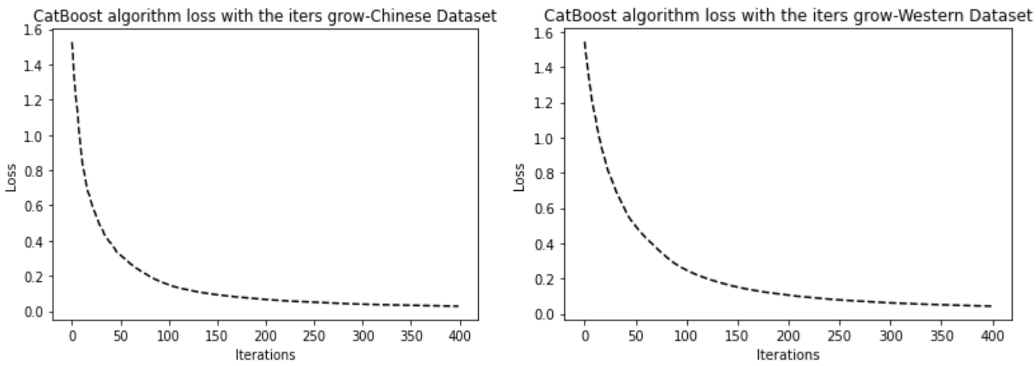


Figure 8.11 CatBoost algorithm loss with the increase in iterations

Table 8.7 The precision, recall, and F-measure of each dataset

Ensemble models												
Dataset	Classification Accuracy											
	Precision				Recall				F-Score			
	RF	XGB	CB	ADA	RF	XGB	CB	ADA	RF	XGB	CB	ADA
Chinese	0.882	0.939	0.943	0.932	0.881	0.938	0.943	0.931	0.881	0.938	0.942	0.931
Western	0.74	0.717	0.743	0.717	0.758	0.764	0.798	0.764	0.7227	0.737	0.766	0.735

8.4 Discussion

The purpose of this chapter is to show that the proposed user profile strategy is effective. To do so, this section discussed how the amins was achieved one point by pone point.

8.4.1 Propose and Evaluate the Recommendation Model

In this study, a DT-based P2P ETRS that proposes 5 platforms to trading users was developed. It utilises a set of features and these five categories are categorised according to relevant energy consumption and trading mediations knowledge, which enhances the classification accuracy rate and decreases the DT’s complexity level. The classification accuracies were recorded at 86% and 64% in the Chinese and Western datasets, respectively. In the Chinese dataset, a classification accuracy rate of 94% was achieved for the retailing classification, and 72% for the blockchain classification. Conversely, in the Western dataset, a classification accuracy rate of 86% was obtained for the blockchain classification and 48% for the hybrid classification.

To test whether the ensemble learning tree models increase the accuracy, the second experiment was conducted on the following four classifiers: (i) random forest, (ii) AdaBoost, (iii) XGBoost, and (iv) CatBoost. The primary finding is that bagging performed significantly weaker than boosting and the base learner. The results indicated that there were no statistically material differences between the boosting method and the individual models. Boosting classifiers demonstrated a marked improvement over the base learner, decision tree. The boosting methods' superior performance to the DT classifiers can be attributed to DTs sensitivity to perturbation on the training samples, in addition to their status as unstable learners. Furthermore, the results found CatBoost to be the most effective classifier for enhancing from the baseline classifier, DT, whereas DT was the weakest classifier.

In general, the classification accuracy is improved by using ensemble models. Based on the summary of the results, if the primary focus is the overall performance of the models in predicting platforms, CatBoost achieved the highest results using significant features. The performance of ensemble models is also compared using all features and using important features as input. Tree-based ensembles showed better performance to predict platform choice due to their ability to learn non-linear solutions and these models can be scaled on large datasets; boosting the algorithm corrects previous mistakes by the model and subsequently improves the performance by rectifying previous results and enhancing the performance (Bauer & Kohavi, 1999; Zhang & Haghani, 2015).

8.4.2 Reducing the Amount of User Input

Reducing the user input is important for the user profile construction strategy, as it can mitigate the risk of user information exposure and user information exiting in the primary entry. Thus, feature selection is the best solution to reduce user input.

There was no significant meaning in improving recommendation accuracy to limit the factors in the Chinese dataset shown in Figure 8.5. The experimental results from 8.5a indicate that combining consumption value factors improves the recommendation accuracy on the Chinese dataset, however it does not matter as much in the Western dataset (8.5b). This suggests that the Chinese database has a more evenly distributed influence on final product choice across features on trading behaviour. Still, the Western database, with features on consumption value, has a more dispersed influence.

Regarding the performance of the feature-selection algorithms, the mRMR algorithm is useful for the western dataset but not for the Chinese dataset. Experiment one identified only 5 features which are important to the western model within mRMR. Meanwhile, 20 significant features in the Chinese dataset were selected, almost half of all of the available features, as the number of the features cannot influence the accuracy of the model.

DT has its own advantages on interpretation, however there is a risk of overfitting (Bramer, 2007). Significant features that are a subset of the original feature set are identified by DT. DT, combined with significant features, not only improves prediction performance but also reduces the cost of the data collection. There are 42 features regarding the trading performance dataset to measure the predictive platform choice. Considering 5 important variables identified by DT significantly improves the prediction process of platform selection on the western dataset.

8.4.3 Analysing P2P Energy Trading Products Selection Behaviour with TCV

Figures 8.12 and 8.13 demonstrate the importance of features in the Chinese and Western datasets. Feature importance decides features are finally selected for the RS; in other words, extracted factors are important to the system. Figure 8.12 suggests the five most influential factors for Chinese energy trading users are their feeling of uncertainty towards energy trading products, house status, income level, service quality, and economic benefits.

Feeling of uncertainty

As evident in Figure 8.12, feeling of security is the most important feature for Chinese participants. This means Chinese participants are more influenced by emotion values when they select the P2P energy product. The feeling of security can decide the trust development of consumers; trust can further influence a consumer's selection of technological products or services (Kesharwani & Bisht, 2012). This finding aligns with previous literature that emotion can influence consumer behaviour, such as product selection (He & Mukherjee, 2007; Khan & Mohsin, 2017; Liu et al., 2021; Mazaheri et al., 2011; Quintal et al., 2010).

House status

The results suggest that house status can influence Chinese consumers decision when selecting energy trading products. House status decides to what extent consumers can alter their house and the price of the installation and usage. This finding aligns with previous literature that building information can affect the economic consideration of the adoption of a renewable energy products and further influence consumers' behaviour (Coffman et al., 2018; Sommerfeld et al., 2017)

Income level

As one of the most important demographic aspects in conditional value, income level is a determinant for Chinese consumers to select an energy trading platform. According to the study by Barbose et al. (2018), residents with a higher income level are more likely to adopt green products than those with a lower income level. This finding also aligns with previous literature on how income level influences the selection of various products (Liu & Zhang, 2020; Maamor et al., 2016; Samukange et al., 2020).

Service quality

Figure 8.12 illustrates that Chinese consumers pay more attention to service quality than other dimensions of system benefit, such as information quality and system quality, which aligns with the findings of previous literature. As a service, energy trading platforms must respond promptly when systems fail or customers have problems using those systems (Buchanan & McMenemy, 2012; DeLone & McLean, 1992). The importance of service quality is also consistent with the finding that Chinese consumers' feeling of uncertainty can influence their selection of products as discussed above.

Economic benefits

Economic benefits involve two aspects: the price of using a certain energy trading product and how much money consumers can save using the platform. This finding aligns with the ISS model and classic economic utility theory (DeLone & McLean, 1992; Stigler, 1950). The change in price and money saved leads to the change in net utility. For example, when price increases and other dimensions stay the same, the net utility of using the product decreases. Consumers' behaviour may change according to how they perceived the change in net utility (Buczynski, 2013; Gansky, 2010; Hennig-Thurau et al., 2007; Lamberton & Rose, 2012).

Sustainability consideration

As showed in Figure 8.13, western consumers consider environment sustainability as an important factor when selecting energy trading platforms. One of the most important reasons why people prefer energy trading to traditional energy suppliers are the beneficial impacts of energy trading on protecting the environment (Botsman & Rogers, 2010). Previous study reveals that consumer purchase behaviour involved the sustainably concern and that customers have become more cognizant of picking environmentally friendly products (de Medeiros & Ribeiro, 2017; Panda et al., 2020). This means the sustainability itself is included as part of the intrinsic value of the product and thus, for the western respondents, the fundamental motivation to select the products is sustainability. With more awareness of environment pressure, people tend to choose environment friendly products (Gansky, 2010; Goyal, 2017; Shabbir et al., 2020)

Cultural aspect: individualism-collectivism

The results indicate the individualism or collectivism level of Western consumers can influence their choice of energy trading platforms. This dimension of culture, proposed by (Hofstede, 1980), measures how consumers feel regarding the relationship between themselves and the societies around them. The results show that the collectivism level of Western consumers have an essential impact on their energy trading products selection. Aligning with the previous finding on the environment sustainability, collectivistic consumers are more likely to consider the environment around them and other members in the community (Hui & Triandis, 1986; Singelis et al., 1995).

Source of experience and reasons for consumption between individualistic and collectivist societies is informed differently when collaboratively consuming (Hamari et al., 2016); when individualistic cultures have individual motivation to engage in collaborative buying is personal, collectivist culture seeks to adhere to subjective norms. For example, Ni (2021) reviews collaborative consumption in China and found Chinese customers' decisions to be influenced by their need to associate in society.

Thus, when engaging in collaborative consumption, customers first evaluated if consumption is for social good or individualistic benefits. Unlike such collective purchasing that collectivist customers engage in, for individualistic customers, their decision is not controlled by society, but individual motivation comes due to need to save cost, be sustainable, and obtain what they need (Roy et al., 2018). Although one common trend between individualistic and collectivist consumers regarding collaborative buying is sustainability, uniqueness and novelty seeking, collectivist decisions are influenced by how society perceives uniqueness, sustainability, and uniqueness (Bar & Otterbring, 2021).

Novelty seeking

In Western countries, to what extent an individual seeks novelty is a strong predictor of which energy trading platform they select. Whether it's blockchain or cryptocurrency, the novelty seeking of the new stuff can attract curious customers to participant and this aligns with the findings from the interview as discussed before. Also, as M3 explained in the interview, ‘Some customers are more willing to try energy trading precisely because it can be done in cryptocurrencies’

As previous literature argues, consumers need a certain level of stimulation and this novelty seeking tendency can lead them to try new energy trading platforms (Berlyne, 1960; Howard & Sheth, 1969; Poon & Huang, 2017). This finding also aligns with empirical studies on how novelty seeking affects consumers’ behaviour, in terms of searching, trialling, and switching to new products or brands (Kim et al., 2020; Sugandini et al., 2018; Zeba & Ganguli, 2019; Zhuang et al., 2017).

Reputation

Reputation is an important factor extracted from the topical analysis from online big data and it plays an essential role when Western consumers select energy trading platforms. The reputation of a product can lead to the trust of that product or brand, and the trust can further lead to the change of consumers behaviour (Craciun & Moore, 2019; Fedorko et al., 2017; Reyes-Menendez et al., 2019; Ryan & Casidy, 2018). The reputation of online product, such as energy trading platforms, is even more important. For example, Wang et al. (2020) and Choi and Burnham (2020) confirm the impact of product or service reputation on consumers’ sharing behaviour.

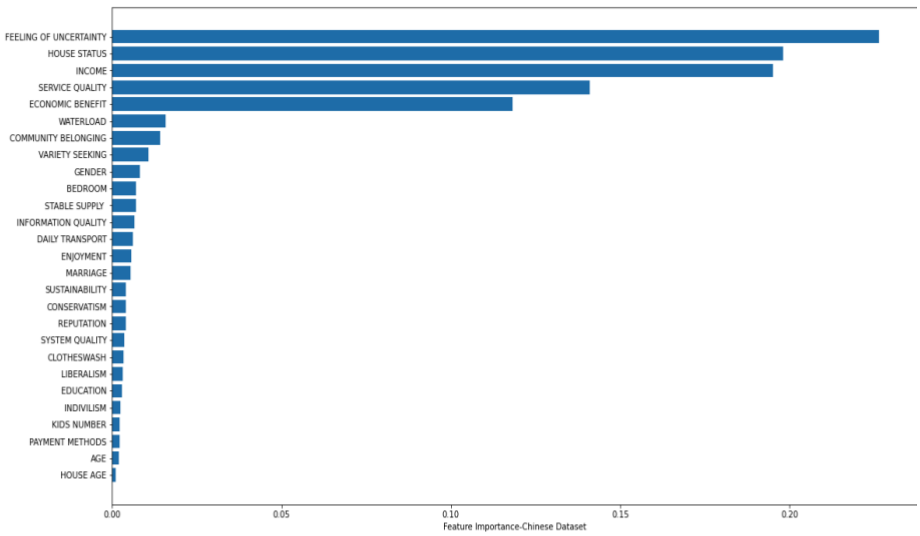


Figure 8.12 Feature importance ranking -Chinese dataset

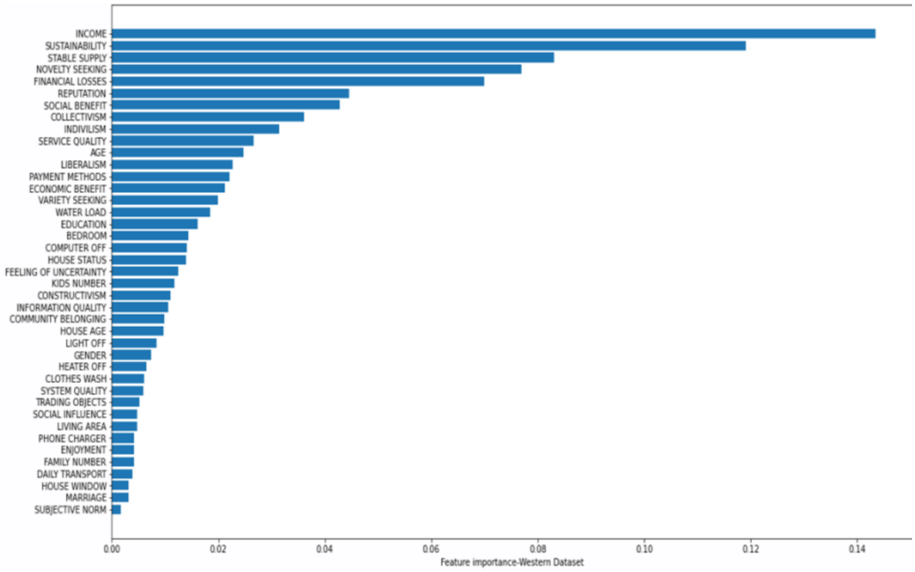


Figure 8.13 Feature importance ranking -Western dataset

The results suggest that income level plays an essential role when Chinese and Western consumers select energy trading products. This means the important role of income level on both regions do not differ because of the different environment of China and Western countries. This finding aligns with previous literature that consumers' demographic information, such as income level, can affect their behaviour on product selection (Barbose et al., 2018; Rai & Robinson, 2015; Sardianou & Genoudi, 2013; Sovacool et al., 2012).

The results also paint the difference between Chinese and Western consumers, in terms of the motivation behind adopting energy trading products. First, the conditional values are mentioned for both groups, but the specific items are totally different. The important conditional value that Chinese participants concern is house information and income. For Western participants, on the other hand, individualism-collectivism is more important. This finding confirms the argument by Sheth et al. (1991) that different contexts or situations can lead to different consumer behaviour.

Second, other than different conditional values, Chinese participants focus more on the feeling of uncertainty than Western participants who are more influenced by epistemic and social values (Jun Li et al., 2019). This difference surprisingly contradicts the findings of previous literature. According to the Hofstede scores, uncertainty is more acceptable in China than in Western countries because uncertainty avoidance score of China (30) is less than the score of Western countries (e.g. Australia: 51; Germany: 65; United Kingdom: 35; United States: 46) (Hofstede, 1980).

Thirdly, building information is important to the P2P products selection in the Chinese Dataset, but the western dataset is not the same. This means that Chinese respondents are more constrained by the condition of the house itself. As showed in Figure 8.12, house status play an important role to the hybrid platform and community platform. That is due to the following two reasons. First, Chinese cities are densely populated, and community-based trading is more responsive to everyone's trading needs. Second, China's residential cities are densely populated meaning there are limited places to install photovoltaic panels. This dictates that the supply of energy is not sufficient. Thus, most respondents still prefer a hybrid platform as the supply of energy is guaranteed.

8.5 Summary

This chapter shows the strength of ensemble learning in predicting preferential trading platform selection to the traders. Most academics still consider this a unique skill; hence it cannot be predicted to show a satisfactory result. Choosing the appropriate classifier for P2P ETRS is a complex undertaking and relies on the data. Thus, DT-based P2P ETRS was investigated, followed by ensemble-based P2P ETRS. Moreover, the effectiveness of the following five classification algorithms were examined and compared in terms of the disparities shown in the datasets: (i) DT, (ii) RF, (iii) AdaBoost, (iv) XGBoost, and (v) CatBoost.

The classification algorithms were assessed with appropriate scientific strategies including classification accuracy rate, confusion matrix, precision, recall, and F-measure scores. Experiment 1 showed that the number of features has diverse impacts in the two datasets. More specifically, in the Chinese dataset, the classification accuracy increases with a higher number of features. Conversely, in the Western dataset, a greater number of features does not enhance its accuracy; therefore, 5 features were ultimately selected for the optimum model. In addition, it was found that the most important features of each dataset differ, meaning that the features that can affect the various platform recommendations are diverse.

Experiment 2 entailed the application of four ensemble methods to build predictive models, including bagging and boosting. Upon comparison of these four ensemble algorithms, all were found to be superior to DT, with CatBoost being the best performer. This confirmed that establishing those ensemble-based algorithms can enhance the classification accuracy of both datasets. For future system improvements in relation to recommendation accuracy and decreasing redundant features, this study recommends utilising ensemble learning methods such as CatBoost.

The interpretation of the prediction models also explains the most influential factors for both Chinese and Western consumers select energy trading products. The results suggest that Chinese energy trading consumers tend to focus on emotional and functional value, while Western consumers consider social and epistemic value more than other dimensions.

Chapter 9: Discussions, Conclusions, Implications, and Limitations

This chapter firstly made a discussion and the plausible reasons that can be used to explain the distinctive findings for this study with comparison with the results found in previous studies (section 9.1). Secondly, this chapter aims to critically evaluate the research questions and objectives outlined in Chapter 1 (section 9.2). Subsequently, a description is provided of the main contributions (section 9.3) and the implications of this research (9.4). Then, the overall limitations and future research direction are made (section 9.5). Finally, the summary of the entire study is included (section 9.6).

9.1 Discussions

This study has adapted a user profile construction strategy to solve the CCSR in the recommendation systems with the theory of consumption values. The following section firstly discussed how this study used user profile construction strategy to solve the CCSR in the P2P ETRS. Furthermore, the second section outlines how the TCV helped to interpret the RS.

9.1.1 CCSR with User Profile Construction Strategy

Cold start problems have been solved with collaborative filtering, contentment-based filtering, hybrid filtering, context awareness, and increasing the data source. Nevertheless, these strategies are not suitable for the CCSR. However, user profile construction is notably effective to solve CCSR.

Features were retrieved from the expert interviews and UGC analysis and then compared to the features listed in earlier studies. Finally, as part of the P2P energy trading user profile dataset, a survey was used to collect data from various countries.

Two experiments are used in this work to examine the efficacy of the suggested user profile construction technique using machine learning. The goal is to assess the decision tree-based classifier's performance and see how an ensemble-based classifier might improve ETRS classification accuracy. In terms of system performance, the accuracy of DT model achieves an 86.1% for the Chinese dataset and a 68.5% for the western data set. In addition, ensemble learning algorithms can improve the classification on both datasets, as CatBoost performed best in the ensemble model, with an accuracy rate of 94.3% on the Chinese dataset and 79.8% on the western dataset.

9.1.2 Recommendations with Theory of Consumption Value

Theoretical perspectives have rarely been taken into consideration when building recommendation systems, especially the completely new system. With the theory foundation, the RS easy to be understandable, which is simple in mathematics but difficult to achieve in terms of human behaviour (Ricci et al., 2015). Thus, it is critical for Information System research to explain human behaviour anticipated by RS.

This approach has demonstrated the theory of consumption values and found five important findings compared with other CC actives. Firstly, in terms of the functional values, P2P energy trading paid more attention on the stable supply and financial loss caused by the fluctional of the cryptocracy. Secondly, for emotional values, P2P paid more attention on the feeling of uncertainty rather than enjoyment. Thirdly, in terms of epistemic values, P2P energy trading paid more attention on novelty seeking rather than variety seeking. Finally, in terms of the conditional values, income of demographic information and building status of the building information played a more important role.

In addition, compared with the two datasets, the most important consumption values are different. Firstly, income is important for both datasets. Secondly, the contents of functional values are different in the two datasets. Chinese users paid more attention on felling of uncertainty while western users paid more attention to sustainability, stable supply, and financial loss. Thirdly, epistemic value is important for the western dataset, but not in the Chinese dataset.

9.2 Revisiting the Research Objectives

Firstly, to review the extant literature on RSs and CSR and identify research challenges and gaps.

This study critical evaluated the different approaches to deal with the RSs and summarised advantages and drawbacks between the approaches, the CB, CF, HR and ML. CBF can easily find the similarities among different items, but it only provides suggestions based on user's current interests (Casey et al., 2008; Liu et al., 2010). While CF makes it difficult to form recommendations to the new items (Nagarnaik & Thomas, 2015; Ramakrishnan et al., 2020), if an item is not seen during training, the system cannot create an embedding for it and cannot use the model to make predictions for this item. ML approaches compass different algorithms and possess advantages and disadvantages of each.

However, P2P energy trading system as a new system faces the CCSR problem. Hybrid algorithms, context awareness, increasing the diversity of the data source, and use profile construction are the major four strategies introduced to mitigate the CSR. This study adapted the user profile construction strategy, because the lack of historical user data deprives the RS of an essential input (Khribi et al., 2015); thus, the CBF, CF, and HR cannot be applied in this new system.

Secondly, to design and develop a survey for user profile construction dataset from different countries

Most of user profile construction strategy are based on the generation of demographic information and personality information to push. As noted in Chapter 3, due to energy consumption features, the selection behaviour for energy trading is likely to be different from those of other P2P activities (e.g., using Airbnb or Uber).

A mix-method approach completed RS by building the trading datasets. An exploratory sequential design approach (Maxwell, 2008) was used to collect sufficient data. The mixed methodology examines the factors shaping energy trading and offers precise, tailored trading recommendations. The mix-method approach includes an expert interview, user-generated content analysis, survey, which comprise the first phase of the study. This is important for the construction of the ETRS user profile dataset. Besides, I collected two different datasets, one Chinese dataset and one Western dataset to investigate the different preferences in the trading.

I extracted the features from the result of the expert's interview analysis and UGC analysis and then compared with the features in previous studies. I finally used survey to collect the data as the P2P energy trading user profile dataset. besides, as part of this study, 1500 questionnaires on human behaviour were carried out by volunteers from the 5 most favoured platform types in various countries.

Thirdly, to evaluate the proposed recommendation model for P2P ETRS by using a variety of techniques.

This study employs machine learning via two experiments to test the effectiveness of the proposed user profile construction strategy. The objective is to examine the performance of the decision tree-based classifier and determine how the ensemble-based classifier can render the ETRS classification more accurate. The two experiments located appropriate features and optimal models from the different algorithm classifications and evaluated the combination method classifications. The rationale for the diverse classifier selection in the model construction appears in the model subsection. Furthermore, the ensemble-based classifier compares the selected classifier results to determine whether elevated classification accuracy is achievable.

Fourthly, to reduce the user input in the user profile construction

Feature selection is the best solution to reduce user profile inputs. In the first experiment, the feature selection algorithm was employed to ascertain whether there are any correlations between the limited features and classification accuracy. The test-classification accuracy was assessed using various numbers of features, and the results confirmed that utilising more features does enhance it in the Chinese dataset.

Fifthly, to compare different consumption values of P2P energy trading by comparing other collaborative consumption actives

The consumption values played on P2P energy trading are different with other CC actives. Regarding to the functional value, P2P energy trading considered the stable supply and the financial loss which is differ with previous motivations on CC. In terms of the emotion value, P2P energy trading considered the feeling of uncertainty while other P2P actives, such as Airbnb considered enjoyment. Besides, in terms of epistemic value different with other CC, the variety seeking do not play a significant role on P2P energy trading, but novelty seeking played an important role in the western dataset. Finally, for the conditional value, house status is important for P2P energy trading rather than any other CC else.

Sixthly, to understand consumers' most important perceived value when selecting energy trading platforms considering their cultural background.

This study uses the results of feature importance and finds that products selection motivations from China and Western countries are different, though income level plays a critical role in both cultures. The most important value for Chinese consumers is emotional and functional value. For Western consumers, however, social, and epistemic value is more important.

Also, the conditional values are mentioned for both groups, but the specific items are totally different. The important conditional value that Chinese participants concern is house information and income. while individualism-collectivism is more important in the western dataset.

In the Chinese dataset, building information is significant for P2P product selection, however this is not the case in the western dataset. This suggests that Chinese respondents are more influenced by the house's condition.

9.3 Contributions

This study is the first study that uses the user profile construction in solving CCSR in the P2P energy trading ETRS. Energy trading products selections and recommendation systems are significant topics in both consumer behaviour and Computer Science. As such, it is impossible to solve the problem in all aspects within one study. Thus, this study focuses on mitigating CCSR on P2P ETRS by adapting user profile construction strategy within the theory of consumption value. Two contributions have been made by addressing the research gaps.

Firstly, the researchers attempt to solve the complete cold start problem using the user profile construction strategy. User profile construction is a common strategy in dealing with the new system recommendation. However, this study used the theory of consumption value to provide theoretical support to build the user profile dataset. After collecting data from the survey, a decision tree-based recommendation model was first proposed. The classification accuracies were recorded at 86% and 64% in the Chinese and Western datasets, respectively in the decision model. However, in the second experiment, I compared with the Random Forest, AdaBoost, XGBoost, CatBoost and found the accuracy was improved to 94.3% under the CatBoost. That is to say, the ensemble learning improved the accuracy.

Moreover, reducing user input contributed to adapting the user profile construction strategy. In Machine Learning, the algorithms are mostly built on well-picked datasets which are difficult to obtain in the energy trading environment. The importance of each feature is ranked in line with the optimal model and the SHAP value to illustrate the different decisions on the optimal model. Thus, the DT-based ETRS and ensemble-based ETRS are compared to investigate the performance of different models. Furthermore, feature selection was applied under the DT model to investigate whether the limited features group, such as the trading behaviour, can help to improve the overall accuracy. On the western dataset with the fewer features, increased accuracy indicates that reducing the number of the features can improve the barrier of user's privacy exposure as the users do not need to fill in so many questions in the following particle stage.

Secondly, a bridge between ML and the traditional social science theory is created. Findings of previous studies have suggested a link between the algorithm and traditional theories. This study demonstrates that the complete cold start problem could be built up with the traditional theories such as the theory of consumption values.

This study confirms the influence of conditional value on consumers' product selection (Sheth et al., 1991). As Sheth et al. (1991) propose, consumers behaviour may change in different contexts and environments. Compared with other CC actives, consumption values used in P2P energy trading differ. P2P energy trading considers a stable supply and a financial loss in the functional value. In the emotion value, P2P energy trading considered the feeling of uncertainty, whereas other CC activities considered the sensation of enjoyment. Furthermore, in terms of epistemic value in comparison to other CC, variety seeking does not play a significant role in P2P energy trading, whereas novelty seeking does in the western dataset. In addition, house status is vital for P2P energy trading.

By comparing the motivations under different cultural backgrounds, the interpretation of the model identifies different essential factors for Chinese and Western consumers when they select energy trading products. Specifically, the feeling of uncertainty and the functional value of the system is critical for Chinese consumers. On the other hand, the relationship between individuals, their community, the environment, and the degree of novelty-seeking influences product selection for Western energy trading platform users.

9.4 Implications of the Study

As a new study this recommendation approach can provide a benchmark for future ETRSs studies. User profile construction strategy was adapted, and the study proves it can effectively solve the CCSR in the P2P ETRS. This strategy can be used in any other recommendation scenarios from scratch, particularly for charity P2P platforms. In other words, this strategy can be applied in the future to make further recommendations. Future researchers could expand the study into other contexts of the new system, such as the IoT system, Smart Homes, and Smart City systems.

Secondly, more machine learning algorithms can be applied if the sample size is large enough, such as deep learning. With the exponential growth of data and information available via the internet, deep learning can play an important role in the RSs (Da'u & Salim, 2020; Wen, 2021). Particularly when the P2P ETRS is running and the system has a small amount of data, deep learning can be combined with traditional methods, such as collaborative filtering and hybrid approach, which might be helpful for the improvement of the recommendation performance.

Thirdly, more theories can be reviewed to fill the research gap. TCV is effective in setting up the new systems. However, TCV only explain the selected behaviour in a certain time; it cannot cover all the users' consumption behaviour in a dynamic period. It is believed that many other traditional theories could be borrowed and applied into the different recommendation system context, such as marketing theories, psychological theories.

Furthermore, practical implications exist in this study. Firstly, security issues during the trading process must be closely monitored. The lack of transparency in this new market creates issues with privacy and security for customers, since many DERs have different owners and their characteristics vary accordingly. Concerning IoT nodes, for example, it would not only be inappropriate for them to trade in unregulated energy markets on a large scale, but this would also be considered unsafe. Moreover, when IoT nodes have surplus energy, their privacy concerns may inhibit them from doing business with energy suppliers.

However, with the help of understanding the decisions made under TCV, accurate recommendations may be offered to new users through the system. It is necessary to recommend the products integrated with the main grids and clearing with local currency. For the western users, as novelty seeking is important, the blockchain products can be recommended. But before recommendation, it is necessary to address the issues of stable supply and brand reputation.

Secondly, the question enquiring stage in the real scenario avoids asking users to enter as much information as possible, reducing the risk of privacy exposure. On the Chinese dataset, the proposed ETRS used relevant and non-redundant inputs of features to achieve the best recommendation results. Nevertheless, features must be extracted from all feature groups, not simply the influence of one group. At the same time, this creates difficulties in practice in getting users to fill in personal information. This is because the risk of information leakage increases along with the amount of personal information available to traders. In practice, however, we can filter the characteristics of new users by selecting only the options with high importance based on the importance of each feature.

9.5 Limitations and Future Research Directions

As discussed above, this study mitigates the CCSR in P2P ETRS. However, several limitations are inevitable.

Firstly, a critical obstacle in RS relates to the inability of the structures to show innovative tendencies. According to Kim and Chen (2015), algorithms can obtain information from product databases, and the use of human operators in website design heightens the likelihood of bias. Nevertheless, in the case of energy consumption, multiple challenges exist in identifying the dynamic preferences associated with energy consumption. In other words, the user profile construction is based on the static input, while the users' preference always changes. Product perception and popularity are constantly changing as new selection emerges. Thus, customer inclinations are evolving, leading them to redefine their taste.

Future study may address this by considering more soft constraint aspects to construct a dynamic dataset. The soft constraints include two aspects. First, dynamic factors such as the time may be considered. Future studies should consider using timeSVD++, which can simulate the temporal dynamics of user interests by changing static biases and latent factors into time-dependent ones. Thus, modelling temporal dynamics should be key when recommender systems or general customer preference models are designed.

Secondly, this study only considered tree-based algorithms, such as decision tree and enshelling learning. Future studies may use more algorithms to explore the classification algorithm. Other traditional classification algorithms, such as KNN, RTree, and RBF could be used as base learners and could benefit from boosting and bagging methods. In addition, ensemble learning methods, such as stacking, random sub-spaces, or pasting could be employed.

Thirdly, the database used in this study is effective as outliers are removed, data has been normalised, and imbalanced classification has been disposed of. However, it is difficult to achieve a perfect database and the presence of noise is inevitable. In future research, when considering scenarios for the use of energy trading recommender systems, researchers need to consider how to detect anomalous data, for example, by using a deep learning autoencoder.

9.6 Summary of the Study

As a result of the rapid growth in P2P energy trading platforms in the market, the categories are becoming increasingly populated with different contents. When selecting their preferred platforms before actual trade, traders can therefore easily be overwhelmed. A user profile construction strategy is adopted in the study to mitigate the CCSR in the P2P energy trading. Four targets were set and the first one is to construct the dataset by using a mix method research approach. Expert interviews and UGC analysis were used to identity the different factors which can impact the P2P energy products selection. Then, data collected from surveys were collected to construct the P2P energy trading user profile datasets.

The second target is to propose and evaluate the recommendation models for the P2P energy trading recommendation platform. A decision tree model was firstly proposed after data pre-processing, including data cleaning and imbalanced classification. After comparing the ensemble learning classification, the ensemble based P2P ETRS, such as CatBoost, performed better than the decision tree based P2P ETRS.

The third target is to select the useful features to reduce user input. To achieve this target, during the pre-processing, unnecessary inputs that are either irrelevant or redundant were eliminated using the proposed two-feature selection method. The experimental results presented in Chapter 8 confirmed that the proposed ETRS used a small number of relevant and non-redundant inputs of features to achieve the best recommendation results on the western dataset. This means that the proposed system is considered non-intrusive and more likely to be accepted by traders.

Finally, the interpretation of the forecasting model helps to explain the critical factors which can influence consumers' energy trading platforms. This finding contributes to theory and practical implication by suggesting the important information energy trading platforms should obtain from new users for a more accurate recommendation.

References

Abdar, M., & Yen, N. Y. (2020). Analysis of user preference and expectation on shared economy platform: An examination of correlation between points of interest on Airbnb. *Computers in Human Behavior, 107*, 105730.

Abel, F., Herder, E., Houben, G.-J., Henze, N., & Krause, D. (2013). Cross-system user modeling and personalization on the social web. *User modeling and user-adapted interaction*, *23*(2), 169-209.

Abramova, S., & Böhme, R. (2016). Perceived benefit and risk as multidimensional determinants of bitcoin use: A quantitative exploratory study.

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, *17*(6), 734-749.

Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. In *Recommender systems handbook* (pp. 217-253). Springer.

Aitzhan, N. Z., & Svetinovic, D. (2016). Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Transactions on Dependable and Secure Computing*, *15*(5), 840-852.

Aitzhan, N. Z., Svetinovic, D. J. I. T. o. D., & Computing, S. (2018). Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *15*(5), 840-852.

Ajesh, A., Nair, J., & Jijin, P. (2016). A random forest approach for rating-based recommender system. 2016 International conference on advances in computing, communications and informatics (ICACCI),

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Ajzen, I. (2005). *Attitudes, personality, and behavior*. McGraw-Hill Education (UK).

Ajzen, I., & Driver, B. L. (1992). Application of the theory of planned behavior to leisure choice. *Journal of leisure research*, *24*(3), 207-224.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological bulletin*, *84*(5), 888.

Al-Hassan, M., Lu, H., & Lu, J. (2015). A semantic enhanced hybrid recommendation approach: A case study of e-Government tourism service recommendation system. *Decision support systems*, *72*, 97-109.

Alam, M. R., St-Hilaire, M., & Kunz, T. (2017). An optimal P2P energy trading model for smart homes in the smart grid. *Energy Efficiency*, *10*(6), 1475-1493.

Alam, M. R., St-Hilaire, M., & Kunz, T. (2019). Peer-to-peer energy trading among smart homes. *Applied Energy*, *238*, 1434-1443.

Albinsson, P. A., Perera, B. Y., Nafees, L., & Burman, B. (2019). Collaborative consumption usage in the US and India: An exploratory study. *Journal of Marketing Theory and Practice*, *27*(4), 390-412.

Aleksandrova, M., Brun, A., Boyer, A., & Chertov, O. (2017). Identifying representative users in matrix factorization-based recommender systems:

application to solving the content-less new item cold-start problem. *Journal of Intelligent Information Systems*, *48*(2), 365-397.

Ali, J., Khan, R., Ahmad, N., & Maqsood, I. (2012). Random forests and decision trees. *International Journal of Computer Science Issues (IJCSI)*, *9*(5), 272-278.

Alirhayim, Y., Mokdad, I., Vorobyev, D., Gabriel, P., & dos Santos, R. P. (2017). The Ethereum-powered Reputation Platform for Commerce.

Amat-Lefort, N., Marimon, F., & Mas-Machuca, M. (2020). Towards a new model to understand quality in collaborative consumption services. *Journal of Cleaner Production*, *266*, 121855.

Amatulli, C., De Angelis, M., Pino, G., & Jain, S. (2020). Consumer reactions to unsustainable luxury: a cross-country analysis. *International Marketing Review*, *38*(2), 412-452.

Andrawis, R. R., Atiya, A. F., & El-Shishiny, H. (2011). Forecast combinations of computational intelligence and linear models for the NN5 time series forecasting competition. *International Journal of Forecasting*, *27*(3), 672-688.

Aritoni, O., & Negru, V. (2011). A multi-agent recommendation system for energy efficiency improvement. International Conference on e-Technologies and Networks for Development,

Assaker, G., Vinzi, V. E., & O'Connor, P. (2011). Examining the effect of novelty seeking, satisfaction, and destination image on tourists' return pattern: A two factor, non-linear latent growth model. *Tourism Management*, *32*(4), 890-901.

Ávila, J., Riofrio, X., Palacio-Baus, K., Espinoza-Mejia, M., & Saquicela, V. (2015). Semantic recommender systems for digital tv: From demographic stereotyping to personalized recommendations. 2015 Asia-Pacific Conference on Computer Aided System Engineering,

Ayaki, T., Yanagimoto, H., & Yoshioka, M. (2017). Recommendation from access logs with ensemble learning. *Artificial Life and Robotics*, *22*(2), 163-167.

Baehrens, D., Schroeter, T., Harmeling, S., Kawanabe, M., Hansen, K., & Müller, K.-R. (2010). How to explain individual classification decisions. *The Journal of Machine Learning Research*, *11*, 1803-1831.

Bai, S., Zhu, T., & Cheng, L. (2012). Big-five personality prediction based on user behaviors at social network sites. *arXiv preprint arXiv:1204.4809*.

Bannister, D., & Fransella, F. (2019). *Inquiring man: The psychology of personal constructs*. Routledge.

Banouar, O., & Raghay, S. (2016). User profile construction for personalized access to multiple data sources through matrix completion method. *International Journal of Computer Science and Network Security (IJCSNS)*, *16*(10), 51.

Bar, A., & Otterbring, T. (2021). The role of culture and personality traits in board game habits and attitudes: Cross-cultural comparison between Denmark, Germany, and USA. *Journal of Retailing and Consumer Services*, *61*, 102506.

Barbose, G. L., Darghouth, N. R., Hoen, B., & Wiser, R. H. (2018). *Income Trends of Residential PV Adopters: An analysis of household-level income estimates [PowerPoint presentation]*.

Bardhi, F., & Eckhardt, G. M. (2012). Access-based consumption: The case of car

sharing. *Journal of consumer research*, *39*(4), 881-898.
Barkan, O., Koenigstein, N., Yogev, E., & Katz, O. (2019). CB2CF: a neural multiview content-to-collaborative filtering model for completely cold item recommendations. Proceedings of the 13th ACM Conference on Recommender Systems,
Basu, C., Hirsh, H., & Cohen, W. (1998). Recommendation as classification: Using social and content-based information in recommendation. Aaai/iaai,
Batmaz, Z., Yurekli, A., Bilge, A., & Kaleli, C. (2019). A review on deep learning for recommender systems: challenges and remedies. *Artificial Intelligence Review*, *52*(1), 1-37.
Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine Learning*, *36*(1), 105-139.
Bearden, W. O., Netemeyer, R. G., & Teel, J. E. (1989). Measurement of consumer susceptibility to interpersonal influence. *Journal of consumer research*, *15*(4), 473-481.
Bekkerman, R., & Allan, J. (2003). Using Bigrams in Text Categorization.
Belaïd, F. (2017). Untangling the complexity of the direct and indirect determinants of the residential energy consumption in France: Quantitative analysis using a structural equation modeling approach. *Energy Policy*, *110*, 246-256.
Belk, R. (2014). You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research*, *67*(8), 1595-1600.
Belk, R. W. (1974). Application and analysis of the behavioral differential inventory for assessing situational effects in buyer behavior. *ACR North American Advances*.
Benoit, S., Baker, T., Bolton, R., Gruber, T., & Kandampully, J. (2017). A triadic framework for collaborative consumption (CC): Motives, activities and resources & capabilities of actors. *Journal of Business Research*, *79*, 219-227.
Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937-1967.
Bergsma, S., & Van Durme, B. (2013). Using conceptual class attributes to characterize social media users. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers),
Berlyne, D. E. (1960). Conflict, arousal, and curiosity.
Bernardi, L., Kamps, J., Kiseleva, J., & Müller, M. J. (2015). The continuous cold start problem in e-commerce recommender systems. *arXiv preprint arXiv:1508.01177*.
Bhagavatula, C., Feldman, S., Power, R., & Ammar, W. (2018). Content-based citation recommendation. *arXiv preprint arXiv:1802.08301*.
Bhakkad, A., Dharamadhikari, S. C., & Kulkarni, P. A. (2013). Efficient Approach to find Bigram Frequency in Text Document using E-VSM. *International Journal of Computer Applications*, *68*, 9-11.
Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*(2), 197-227.
Binninger, A.-S., Ourahmoune, N., & Robert, I. (2015). Collaborative Consumption And Sustainability: A Discursive Analysis Of Consumer Representations And

Collaborative Website Narratives. *Journal of Applied Business Research (JABR)*, *31*(3), 969-986.

Birtolo, C., & Ronca, D. (2013). Advances in clustering collaborative filtering by means of fuzzy C-means and trust. *Expert systems with applications*, *40*(17), 6997-7009.

Blanco-Fernández, Y., López-Nores, M., Gil-Solla, A., Ramos-Cabrer, M., & Pazos-Arias, J. J. (2011). Exploring synergies between content-based filtering and spreading activation techniques in knowledge-based recommender systems. *Information Sciences*, *181*(21), 4823-4846.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022.

Block, C., Neumann, D., & Weinhardt, C. (2008). A market mechanism for energy allocation in micro-chp grids. Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008),

Bobadilla, J., Ortega, F., Hernando, A., & Bernal, J. (2012). A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-Based Systems*, *26*, 225-238.

Bock, G.-W., Zmud, R. W., Kim, Y.-G., & Lee, J.-N. (2005). Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS quarterly*, 87-111.

Borràs, J., Moreno, A., & Valls, A. (2014). Intelligent tourism recommender systems: A survey. *Expert systems with applications*, *41*(16), 7370-7389.

Botsman, R., & Capelin, L. (2016). *Airbnb: Building A Revolutionary Travel Company. Said Business School Case*. Oxford University

[Record #493 is using a reference type undefined in this output style.]

Bowman, S., & Willis, C. (2003). We media. *How audiences are shaping the future of news and information*, *66*.

Bramer, M. (2007). Avoiding overfitting of decision trees. *Principles of data mining*, 119-134.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123-140.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5-32.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). Classification and Regression Trees, Pacific Grove: Wadsworth. *BreimanClassification and Regression Trees1984*.

Brooks, J., McCluskey, S., Turley, E., & King, N. (2015). The utility of template analysis in qualitative psychology research. *Qualitative research in psychology*, *12*(2), 202-222.

Browne, M. W. (2000). Cross-validation methods. *Journal of mathematical psychology*, *44*(1), 108-132.

Bruno, B., & Faggini, M. (2017). Sharing economy: for an economic taxonomy. *International Journal of Economics and Finance*, *9*(6), 174-178.

Bryman, A. (2016). *Social Research Methods*. Oxford university press.

Buchanan, S., & McMenemy, D. (2012). Digital service analysis and design: The role of process modelling. *International journal of information management*, *32*(3),

251-256.
Buckland, M., & Gey, F. (1994). The relationship between Recall and Precision. *Journal of the American Society for Information Science*, *45*(1), 12-19.
Buczynski, B. (2013). *Sharing is good: How to save money, time and resources through collaborative consumption*. New Society Publishers.
Buel, R. (1972). *Securing the Revolution: Ideology in American Politics*. (Cornell University Press)
Bunjaku, F., Gjorgieva-Trajkovska, O., & Miteva-Kacarski, E. (2017). Cryptocurrencies–advantages and disadvantages. *Journal of Economics*, *2*(1), 31-39.
Cai, G., Lee, K., & Lee, I. (2018). Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos. *Expert systems with applications*, *94*, 32-40.
Cai, X., Hu, Z., Zhao, P., Zhang, W., & Chen, J. (2020). A hybrid recommendation system with many-objective evolutionary algorithm. *Expert systems with applications*, *159*, 113648.
Caldwell, M., Elliot, S., Henry, P., & O'Connor, M. (2020). The impact of political ideology on consumer perceptions of their rights and responsibilities in the sharing economy. *European Journal of Marketing*.
Cali, U., & Fifield, A. (2019). Towards the decentralized revolution in energy systems using blockchain technology. *Int. J. Smart Grid Clean Energy*, *8*(3), 245-256.
Camacho, L. A. G., & Alves-Souza, S. N. (2018). Social network data to alleviate cold-start in recommender system: A systematic review. *Information Processing & Management*, *54*(4), 529-544.
Carfagna, L. B., Dubois, E. A., Fitzmaurice, C., Ouimette, M. Y., Schor, J. B., Willis, M., & Laidley, T. (2014). An emerging eco-habitus: The reconfiguration of high cultural capital practices among ethical consumers. *Journal of Consumer Culture*, *14*(2), 158-178.
Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. Proceedings of the 23rd international conference on Machine learning,
Casey, M. A., Veltkamp, R., Goto, M., Leman, M., Rhodes, C., & Slaney, M. (2008). Content-based music information retrieval: Current directions and future challenges. *Proceedings of the IEEE*, *96*(4), 668-696.
Celata, F., Hendrickson, C. Y., & Sanna, V. S. (2017). The sharing economy as community marketplace? Trust, reciprocity and belonging in peer-to-peer accommodation platforms. *Cambridge Journal of Regions, Economy and Society*, *10*(2), 349-363.
Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, *41*(3), 1-58.
Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, *40*(1), 16-28.
Chasin, F., von Hoffen, M., Cramer, M., & Matzner, M. (2018). Peer-to-peer sharing and collaborative consumption platforms: a taxonomy and a reproducible

analysis. *Information Systems and e-Business Management*, *16*(2), 293-325.
Chatterjee, S. (2019). Explaining customer ratings and recommendations by combining qualitative and quantitative user generated contents. *Decision support systems*, *119*, 14-22.
Chaudhary, R., Jindal, A., Aujla, G. S., Aggarwal, S., Kumar, N., & Choo, K.-K. R. (2019). BEST: Blockchain-based secure energy trading in SDN-enabled intelligent transportation system. *Computers & Security*, *85*, 288-299.
Chen, J. V., Rungruengsamrit, D., Rajkumar, T., & Yen, D. C. (2013). Success of electronic commerce Web sites: A comparative study in two countries. *Information & management*, *50*(6), 344-355.
Chen, S.-C., & Lin, C.-P. (2019). Understanding the effect of social media marketing activities: The mediation of social identification, perceived value, and satisfaction. *Technological Forecasting and Social Change*, *140*, 22-32.
Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining,
Cheng, R., & Tang, B. (2016). A music recommendation system based on acoustic features and user personalities. Pacific-Asia conference on knowledge discovery and data mining,
Cho, H., & Lee, J. S. (2015). The influence of self-efficacy, subjective norms, and risk perception on behavioral intentions related to the H1N1 flu pandemic: A comparison between K orea and the US. *Asian Journal of Social Psychology*, *18*(4), 311-324.
Cho, Y. H., Kim, J. K., & Kim, S. H. (2002). A personalized recommender system based on web usage mining and decision tree induction. *Expert systems with applications*, *23*(3), 329-342.
Choe, J. Y. J., & Kim, S. S. (2018). Effects of tourists' local food consumption value on attitude, food destination image, and behavioral intention. *International Journal of Hospitality Management*, *71*, 1-10.
Choi, L., & Burnham, T. (2020). Brand reputation and customer voluntary sharing behavior: the intervening roles of self-expressive brand perceptions and status seeking. *Journal of Product & Brand Management*.
Choi, T.-M., & He, Y. (2019). Peer-to-peer collaborative consumption for fashion products in the sharing economy: Platform operations. *Transportation Research Part E: Logistics and Transportation Review*, *126*, 49-65.
Chong, A. Y.-L., Ooi, K.-B., & Sohal, A. (2009). The relationship between supply chain factors and adoption of e-collaboration tools: an empirical examination. *International Journal of Production Economics*, *122*(1), 150-160.
Chwieduk, D. A. (2009). Recommendation on modelling of solar energy incident on a building envelope. *Renewable energy*, *34*(3), 736-741.
Clark, C. F., Kotchen, M. J., & Moore, M. R. (2003). Internal and external influences on pro-environmental behavior: Participation in a green electricity program. *Journal of environmental psychology*, *23*(3), 237-246.
Coffman, M., Allen, S., & Wee, S. (2018). *Determinants of Residential Solar*

Photovoltaic Adoption.
Cohen, D., Aharon, M., Koren, Y., Somekh, O., & Nissim, R. (2017). Expediting exploration by attribute-to-feature mapping for cold-start recommendations. Proceedings of the Eleventh ACM Conference on Recommender Systems,
Collis, J., & Hussey, R. (2013). *Business research: A practical guide for undergraduate and postgraduate students*. Macmillan International Higher Education.
Craciun, G., & Moore, K. (2019). Credibility of negative online product reviews: Reviewer gender, reputation and emotion effects. *Computers in Human Behavior*, *97*, 104-115.
Cresswell, J., & Plano Clark, V. (2011). Designing and conducting mixed method research. 2nd Sage. *Thousand Oaks, CA*, *201*.
Crockett, D., & Wallendorf, M. (2004). The role of normative political ideology in consumer behavior. *Journal of Consumer Research*, *31*(3), 511-528.
Cronin Jr, J. J., & Taylor, S. A. (1992). Measuring service quality: a reexamination and extension. *Journal of marketing*, *56*(3), 55-68.
Cropanzano, R., & Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of management*, *31*(6), 874-900.
Crowther, D., & Lancaster, G. (2012). *Research methods*. Routledge.
Cui, C., Wu, T., Hu, M., Weir, J. D., & Li, X. (2016). Short-term building energy model recommendation system: A meta-learning approach. *Applied Energy*, *172*, 251-263.
Cui, S., Wang, Y.-W., & Xiao, J.-W. (2019). Peer-to-peer energy sharing among smart energy buildings by distributed transaction. *IEEE Transactions on Smart Grid*, *10*(6), 6491-6501.
Da'u, A., & Salim, N. (2020). Recommendation system based on deep learning methods: a systematic review and new directions. *Artificial Intelligence Review*, *53*(4), 2709-2748.
Darby, S. J. (2013). Load management at home: advantages and drawbacks of some 'active demand side'options. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, *227*(1), 9-17.
Datta, A., Sen, S., & Zick, Y. (2016). Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. 2016 IEEE symposium on security and privacy (SP),
Daugherty, T., Eastin, M. S., & Bright, L. (2008). Exploring consumer motivations for creating user-generated content. *Journal of interactive advertising*, *8*(2), 16-25.
De Amo, S., Diallo, M. S., Diop, C. T., Giacometti, A., Li, D., & Soulet, A. (2015). Contextual preference mining for user profile construction. *Information Systems*, *49*, 182-199.
de Medeiros, J. F., Marcon, A., Ribeiro, J. L. D., Quist, J., & D'Agostin, A. (2021). Consumer emotions and collaborative consumption: The effect of COVID-19 on the adoption of use-oriented product-service systems. *Sustainable Production and Consumption*, *27*, 1569-1588.
De Medio, C., Limongelli, C., Sciarrone, F., & Temperini, M. (2020). MoodleREC: A recommendation system for creating courses using the moodle e-learning

platform. *Computers in Human Behavior*, *104*, 106168.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, *41*(6), 391-407.

Deldjoo, Y., Dacrema, M. F., Constantin, M. G., Eghbal-Zadeh, H., Cereda, S., Schedl, M., Ionescu, B., & Cremonesi, P. (2019). Movie genome: alleviating new item cold start in movie recommendation. *User modeling and user-adapted interaction*, *29*(2), 291-343.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information systems research*, *3*(1), 60-95.

DeLone, W. H., & McLean, E. R. (2004). Measuring e-commerce success: Applying the DeLone & McLean information systems success model. *International Journal of Electronic Commerce*, *9*(1), 31-47.

Delzendeh, E., Wu, S., Lee, A., & Zhou, Y. (2017). The impact of occupants' behaviours on building energy analysis: A research review. *Renewable and Sustainable Energy Reviews*, *80*, 1061-1071.

Deng, S., Huang, L., & Xu, G. (2014). Social network-based service recommendation with trust enhancement. *Expert systems with applications*, *41*(18), 8075-8084.

Deshpande, M., & Karypis, G. (2004). Item-based top-n recommendation algorithms. *ACM Transactions on Information Systems (TOIS)*, *22*(1), 143-177.

Desmond, M. (2004). Methodological challenges posed in studying an elite in the field. *Area*, *36*(3), 262-269.

Devaraj, S., Easley, R. F., & Crant, J. M. (2008). Research note—how does personality matter? Relating the five-factor model to technology acceptance and use. *Information systems research*, *19*(1), 93-105.

Dhanda, M., & Verma, V. (2016). Recommender system for academic literature with incremental dataset. *Procedia Computer Science*, *89*, 483-491.

Diekmann, A., Jann, B., Przepiorka, W., & Wehrli, S. (2014). Reputation formation and the evolution of cooperation in anonymous online markets. *American sociological review*, *79*(1), 65-85.

Dietterich, T. G. (2000). Ensemble methods in machine learning. International workshop on multiple classifier systems,

Dillahunt, T. R., & Malone, A. R. (2015). The promise of the sharing economy among disadvantaged communities. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems,

Doney, P. M., & Cannon, J. P. (1997). An examination of the nature of trust in buyer–seller relationships. *Journal of marketing*, *61*(2), 35-51.

Dong, X., Li, H., Liu, S., Cai, C., & Fan, X. (2018). How does material possession love influence sustainable consumption behavior towards the durable products? *Journal of Cleaner Production*, *198*, 389-400.

Duan, R., & Deconinck, G. (2010). Future electricity market interoperability of a multi-agent model of the Smart Grid. 2010 International Conference on Networking, Sensing and Control (ICNSC),

Duhachek, A., Tormala, Z. L., & Han, D. (2014). Stability vs. change: The effect of political ideology on product preference. *ACR North American Advances*.

Ecker, F., Hahnel, U. J., & Spada, H. (2017). Promoting decentralized sustainable energy systems in different supply scenarios: the role of autarky aspiration. *Frontiers in Energy Research*, *5*, 14.

Ecker, F., Spada, H., & Hahnel, U. J. (2018). Independence without control: Autarky outperforms autonomy benefits in the adoption of private energy storage systems. *Energy Policy*, *122*, 214-228.

Eickhoff, M., & Neuss, N. (2017). Topic modelling methodology: its use in information systems and other managerial disciplines.

Emerson, R. M. (1976). Social exchange theory. *Annual review of sociology*, *2*(1), 335-362.

Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tourism Management*, *55*, 62-73.

Ertz, M., Durif, F., & Arcand, M. (2017). An analysis of the origins of collaborative consumption and its implications for marketing. *Academy of Marketing Studies Journal*, *21*(1), 1-17.

Esmat, A., de Vos, M., Ghiassi-Farrokhfal, Y., Palensky, P., & Epema, D. (2021). A novel decentralized platform for peer-to-peer energy trading market with blockchain technology. *Applied Energy*, *282*, 116123.

European Commission. (2015). *Study on the effective integration of Distributed Energy Resources for providing flexibility to the electricity system*.

Featherman, M. S., & Pavlou, P. A. (2003). Predicting e-services adoption: a perceived risk facets perspective. *International Journal of Human-Computer Studies*, *59*(4), 451-474.

Fedorko, R., Bacik, R., & Kerulova, V. (2017). The analysis on the importance of the reputation management in relation to e-commerce subjects. *Polish Journal of Management Studies*, *15*.

Fell, M. J. (2019). Social impacts of peer-to-peer energy trading: a rapid realist review protocol. *SocArXiv. February*, *4*.

Feng, J., Xia, Z., Feng, X., & Peng, J. (2021). RBPR: A hybrid model for the new user cold start problem in recommender systems. *Knowledge-Based Systems*, *214*, 106732.

Feng, Q., He, D., Zeadally, S., Khan, M. K., & Kumar, N. (2019). A survey on privacy protection in blockchain system. *Journal of Network and Computer Applications*, *126*, 45-58.

Fernández, A., Garcia, S., Herrera, F., & Chawla, N. V. (2018). SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of artificial intelligence research*, *61*, 863-905.

Fernández-Tobías, I., Braunhofer, M., Elahi, M., Ricci, F., & Cantador, I. (2016). Alleviating the new user problem in collaborative filtering by exploiting personality information. *User modeling and user-adapted interaction*, *26*(2), 221-255.

Fernández-Tobías, I., Cantador, I., Kaminskas, M., & Ricci, F. (2012). *Cross-domain recommender systems: A survey of the state of the art* In Spanish conference on

information retrieval,
Fernández-Tobías, I., Cantador, I., Tomeo, P., Anelli, V. W., & Di Noia, T. (2019). Addressing the user cold start with cross-domain collaborative filtering: exploiting item metadata in matrix factorization. *User modeling and user-adapted interaction*, *29*(2), 443-486.
Ferov, M., & Modrý, M. (2016). Enhancing lambdamart using oblivious trees. *arXiv preprint arXiv:1609.05610*.
Forouzandeh, S., Berahmand, K., & Rostami, M. (2021). Presentation of a recommender system with ensemble learning and graph embedding: a case on MovieLens. *Multimedia Tools and Applications*, *80*(5), 7805-7832.
Fowler Jr, F. J., & Fowler, F. J. (1995). *Improving survey questions: Design and evaluation*. Sage.
Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, *55*(1), 119-139.
Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.
Fu, W., Peng, Z., Wang, S., Xu, Y., & Li, J. (2019). Deeply fusing reviews and contents for cold start users in cross-domain recommendation systems. Proceedings of the AAAI Conference on Artificial Intelligence,
Galbreth, M. R., Ghosh, B., & Shor, M. (2012). Social sharing of information goods: Implications for pricing and profits. *Marketing Science*, *31*(4), 603-620.
Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International journal of information management*, *35*(2), 137-144.
Gansky, L. (2010). *The mesh: Why the future of business is sharing*. Penguin.
Gantner, Z., Drumond, L., Freudenthaler, C., Rendle, S., & Schmidt-Thieme, L. (2010). Learning attribute-to-feature mappings for cold-start recommendations. 2010 IEEE International Conference on Data Mining,
García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2016). Big data preprocessing: methods and prospects. *Big Data Analytics*, *1*(1), 9-31.
Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, *31*(14), 2225-2236.
Golbandi, N., Koren, Y., & Lempel, R. (2011). Adaptive bootstrapping of recommender systems using decision trees. Proceedings of the fourth ACM international conference on Web search and data mining,
Goldberg, D., Nichols, D., Oki, B. M., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, *35*(12), 61-70.
Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, *24*(1), 44-65.
Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1). MIT press Cambridge.
Google Trends. (2021). *The Distribution of the Interest of Peer-to-Peer Energy Trading Topic Over Time*.

https://trends.google.com/trends/explore?q=%2Fm%2F060b5&date=today%205-y&geo=GB
Gopika, D., & Azhagusundari, B. (2014). An analysis on ensemble methods in classification tasks. *International Journal of Advanced Research in Computer and Communication Engineering*, *3*(7), 7423-7427.
Gössling, S., & Michael Hall, C. (2019). Sharing versus collaborative economy: how to align ICT developments and the SDGs in tourism? *Journal of Sustainable Tourism*, *27*(1), 74-96.
Goyal, A. P. (2017). *Environment-Friendly Products—Adapt Green Now: Green Marketing*. Partridge Publishing.
Goyani, M., & Chaurasiya, N. (2020). A review of movie recommendation system: Limitations, Survey and Challenges. *ELCVIA: Electronic Letters on Computer Vision and Image Analysis*, *19*(3), 0018-0037.
Grbovic, M., & Cheng, H. (2018). Real-time personalization using embeddings for search ranking at airbnb. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining,
Gregory, G. D., Munch, J. M., & Peterson, M. (2002). Attitude functions in consumer research: Comparing value–attitude relations in individualist and collectivist cultures. *Journal of Business Research*, *55*(11), 933-942.
Gretzel, U., Mitsche, N., Hwang, Y.-H., & Fesenmaier, D. R. (2004). Tell me who you are and I will tell you where to go: Use of travel personalities in destination recommendation systems. *Information Technology & Tourism*, *7*(1), 3-12.
Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician*, *61*(2), 139-147.
Guerrero, J., Chapman, A., & Verbic, G. (2017). A study of energy trading in a low-voltage network: Centralised and distributed approaches. 2017 Australasian Universities Power Engineering Conference (AUPEC),
Guntuku, S. C., Roy, S., & Weisi, L. (2015). Personality modeling based image recommendation. International conference on multimedia modeling,
Guo, G., Zhang, J., & Thalmann, D. (2014). Merging trust in collaborative filtering to alleviate data sparsity and cold start. *Knowledge-Based Systems*, *57*, 57-68.
Gupta, S., & Goel, S. (2018). Handling user cold start problem in recommender systems using fuzzy clustering. In *Information and Communication Technology for Sustainable Development* (pp. 143-151). Springer.
Gupta, V., & Lehal, G. S. (2009). A survey of text mining techniques and applications. *Journal of emerging technologies in web intelligence*, *1*(1), 60-76.
Guyader, H. (2018). No one rides for free! Three styles of collaborative consumption. *Journal of Services Marketing*.
Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine Learning research*, *3*(Mar), 1157-1182.
Ha, S., & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of Business Research*, *62*(5), 565-571.
Habibi, M. R., Davidson, A., & Laroche, M. (2017). What managers should know about the sharing economy. *Business Horizons*, *60*(1), 113-121.

Hackbarth, A., & Löbbe, S. (2020). Attitudes, preferences, and intentions of German households concerning participation in peer-to-peer electricity trading. *Energy Policy*, *138*, 111238.

Hagiu, A., & Wright, J. (2015). Multi-sided platforms. *International Journal of Industrial Organization*, *43*, 162-174.

Hague, P. N. (2002). *Market research: a guide to planning, methodology & evaluation*. Kogan Page Publishers.

Hahnel, U. J., Herberz, M., Pena-Bello, A., Parra, D., & Brosch, T. (2020). Becoming prosumer: Revealing trading preferences and decision-making strategies in peer-to-peer energy communities. *Energy Policy*, *137*, 111098.

Halees, A. (2011). Mining opinions in user-generated contents to improve course evaluation. International conference on software engineering and computer systems,

Hamari, J., Sjöklint, M., & Ukkonen, A. (2016). The sharing economy: Why people participate in collaborative consumption. *Journal of the association for information science and technology*, *67*(9), 2047-2059.

Han, C. M. (2017). Individualism, collectivism, and consumer animosity in emerging Asia: Evidence from Korea. *Journal of Consumer Marketing*.

Han, S. C., Chung, H., & Kang, B. H. (2012). It is time to prepare for the future: forecasting social trends. In *Computer applications for database, education, and ubiquitous computing* (pp. 325-331). Springer.

Han, Y., Han, Z., Wu, J., Yu, Y., Gao, S., Hua, D., & Yang, A. (2020). Artificial intelligence recommendation system of cancer rehabilitation scheme based on iot technology. *Ieee Access*, *8*, 44924-44935.

Haneczok, J., & Piskorski, J. (2020). Shallow and deep learning for event relatedness classification. *Information Processing & Management*, *57*(6), 102371.

Hara, S., & Hayashi, K. (2016). Making tree ensembles interpretable. *arXiv preprint arXiv:1606.05390*.

Hartl, B., Hofmann, E., & Kirchler, E. (2016). Do we need rules for “what's mine is yours”? Governance in collaborative consumption communities. *Journal of Business Research*, *69*(8), 2756-2763.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* . Springer Science & Business Media.

Havers, A. (2016). Powerpeers. *Vattenfall*. https://group.vattenfall.com/press-and-media/news--press-releases/newsroom/2016/power-from-your-neighbours

Hawapi, M. W., Sulaiman, Z., Kohar, U. H. A., & Talib, N. A. (2017). Effects of perceived risks, reputation and electronic word of mouth (e-WOM) on collaborative consumption of uber car sharing service. IOP Conference Series: Materials Science and Engineering,

Hayes, B. P., Thakur, S., & Breslin, J. G. (2020). Co-simulation of electricity distribution networks and peer to peer energy trading platforms. *International Journal of Electrical Power & Energy Systems*, *115*, 105419.

Hazrati, N., & Elahi, M. (2021). Addressing the New Item problem in video recommender systems by incorporation of visual features with restricted

Boltzmann machines. *Expert Systems*, *38*(3), e12645.
He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE transactions on knowledge and data engineering*, *21*(9), 1263-1284.
He, H., & Mukherjee, A. (2007). I am, ergo I shop: does store image congruity explain shopping behaviour of Chinese consumers? *Journal of Marketing Management*, *23*(5-6), 443-460.
Hennig-Thurau, T., Henning, V., & Sattler, H. (2007). Consumer file sharing of motion pictures. *Journal of marketing*, *71*(4), 1-18.
Herce-Zelaya, J., Porcel, C., Bernabé-Moreno, J., Tejeda-Lorente, A., & Herrera-Viedma, E. (2020). New technique to alleviate the cold start problem in recommender systems using information from social media and random decision forests. *Information Sciences*, *536*, 156-170.
Hesse-Biber, S. (2010). Qualitative approaches to mixed methods practice. *Qualitative Inquiry*, *16*(6), 455-468.
Hirschman, E. C. (1980). Innovativeness, novelty seeking, and consumer creativity. *Journal of consumer research*, *7*(3), 283-295.
Hoepfl, M. C. (1997). Choosing Qualitative Research: A Primer for Technology Education Researchers. *Journal of Technology Education*, *9*(1), 47-63.
Hofmann, E., Hartl, B., & Penz, E. (2017). Power versus trust–what matters more in collaborative consumption? *Journal of Services Marketing*.
Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, *10*(4), 15-41.
Hollowell, J. C., Rowland, Z., Kliestik, T., Kliestikova, J., & Dengov, V. V. (2019). Customer loyalty in the sharing economy platforms: how digital personal reputation and feedback systems facilitate interaction and trust between strangers. *Journal of Self-Governance and Management Economics*, *7*(1), 13-18.
Holzinger, A., Valdez, A. C., & Ziefle, M. (2016). Towards interactive recommender systems with the doctor-in-the-loop. *Mensch und Computer 2016–Workshopband*.
Hooker, S., Erhan, D., Kindermans, P.-J., & Kim, B. (2018). Evaluating feature importance estimates.
Hou, J., Li, Q., Meng, S., Ni, Z., Chen, Y., & Liu, Y. (2019). DPRF: a differential privacy protection random forest. *Ieee Access*, *7*, 130707-130720.
Howard, J. A., & Sheth, J. N. (1969). The theory of buyer behavior. *New York*, *63*.
Hu, S., Yan, D., Guo, S., Cui, Y., & Dong, B. (2017). A survey on energy consumption and energy usage behavior of households and residential building in urban China. *Energy and buildings*, *148*, 366-378.
Huang, G., Wu, L., Ma, X., Zhang, W., Fan, J., Yu, X., Zeng, W., & Zhou, H. (2019). Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. *Journal of Hydrology*, *574*, 1029-1041.
Huang, L., Li, Y., Huang, X., & Zhou, L. (2021). How social distance affects the intention and behavior of collaborative consumption: A study based on online car-hailing service. *Journal of Retailing and Consumer Services*, *61*, 102534.

Huang, T. C.-K., Chen, Y.-L., & Chen, M.-C. (2016). A novel recommendation model with Google similarity. *Decision support systems*, *89*, 17-27.

Huang, Z., Shan, G., Cheng, J., & Sun, J. (2019). TRec: An efficient recommendation system for hunting passengers with deep neural networks. *Neural Computing and Applications*, *31*(1), 209-222.

Hui, C. H., & Triandis, H. C. (1986). Individualism-collectivism: A study of cross-cultural researchers. *Journal of cross-cultural psychology*, *17*(2), 225-248.

Hull, C. L. (1943). Principles of behavior: An introduction to behavior theory.

Hwang, J., & Griffiths, M. A. (2017). Share more, drive less: Millennials value perception and behavioral intent in using collaborative consumption services. *Journal of Consumer Marketing*.

Hwangbo, H., Kim, Y. S., & Cha, K. J. (2018). Recommendation system development for fashion retail e-commerce. *Electronic Commerce Research and Applications*, *28*, 94-101.

Ianole-Călin, R., Francioni, B., Masili, G., Druică, E., & Goschin, Z. (2020). A cross-cultural analysis of how individualism and collectivism impact collaborative consumption. *Resources, Conservation and Recycling*, *157*, 104762.

Im, I., Kim, Y., & Han, H.-J. (2008). The effects of perceived risk and technology type on users' acceptance of technologies. *Information & management*, *45*(1), 1-9.

IRENA. (2020). *Peer to Peer Electricity Treading Innpvation Landscape Brief*.

Ivashchenko, A. I. (2016). Using cryptocurrency in the activities of Ukrainian small and medium enterprises in order to improve their investment attractiveness.

Jamil, F., Iqbal, N., Ahmad, S., & Kim, D. (2021). Peer-to-peer energy trading mechanism based on blockchain and machine learning for sustainable electrical power supply in smart grid. *Ieee Access*, *9*, 39193-39217.

Jang, S. S., & Feng, R. (2007). Temporal destination revisit intention: The effects of novelty seeking and satisfaction. *Tourism Management*, *28*(2), 580-590.

Jiang, Y., Balaji, M., & Jha, S. (2019). Together we tango: Value facilitation and customer participation in Airbnb. *International Journal of Hospitality Management*, *82*, 169-180.

Jin, X., Zhou, Y., & Mobasher, B. (2005). A maximum entropy web recommendation system: combining collaborative and content features. Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining,

Jing, R., Xie, M. N., Wang, F. X., & Chen, L. X. (2020). Fair P2P energy trading between residential and commercial multi-energy systems enabling integrated demand-side management. *Applied Energy*, *262*, 114551.

Jøsang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision support systems*, *43*(2), 618-644.

Joshi, R., & Alehegn, M. (2017). Analysis and prediction of diabetes diseases using machine learning algorithm: Ensemble approach. *International Research Journal of Engineering and Technology*, *4*(10), 426-435.

Joskow, P. L. (2000). Why do we need electricity retailers?; or, can you get it cheaper wholesale?

Jung, K., Garbarino, E., Briley, D. A., & Wynhausen, J. (2017). Blue and red voices: Effects of political ideology on consumers' complaining and disputing behavior. *Journal of consumer research*, *44*(3), 477-499.

Jurek, A., Bi, Y., Wu, S., & Nugent, C. (2014). A survey of commonly used ensemble-based classification techniques. *The Knowledge Engineering Review*, *29*(5), 551.

Kacen, J. J., & Lee, J. A. (2002). The influence of culture on consumer impulsive buying behavior. *Journal of consumer psychology*, *12*(2), 163-176.

Kankanhalli, A., Tan, B. C., & Wei, K.-K. (2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS quarterly*, 113-143.

Karakayali, N., Kostem, B., & Galip, I. (2018). Recommendation systems as technologies of the self: Algorithmic control and the formation of music taste. *Theory, Culture & Society*, *35*(2), 3-24.

Karami, M., & Madlener, R. (2022). Business models for peer-to-peer energy trading in Germany based on households' beliefs and preferences. *Applied Energy*, *306*, 118053.

Katarya, R., & Verma, O. P. (2016). A collaborative recommender system enhanced with particle swarm optimization technique. *Multimedia Tools and Applications*, *75*(15), 9225-9239.

Kaur, P., Dhir, A., Talwar, S., & Ghuman, K. (2021). The value proposition of food delivery apps from the perspective of theory of consumption value. *International Journal of Contemporary Hospitality Management*.

Kearns, M., & Valiant, L. (1994). Cryptographic limitations on learning Boolean formulae and finite automata. *Journal of the ACM (JACM)*, *41*(1), 67-95.

Kesharwani, A., & Bisht, S. S. (2012). The impact of trust and perceived risk on internet banking adoption in India: An extension of technology acceptance model. *International journal of bank marketing*.

Khan, I., Mahmood, A., Javaid, N., Razzaq, S., Khan, R., & Ilahi, M. J. a. p. a. (2013). Home energy management systems in future smart grids.

Khan, S. N., & Mohsin, M. (2017). The power of emotional value: Exploring the effects of values on green product consumer choice behavior. *Journal of Cleaner Production*, *150*, 65-74.

Khribi, M. K., Jemni, M., & Nasraoui, O. (2015). Recommendation systems for personalized technology-enhanced learning. In *Ubiquitous learning environments and technologies* (pp. 159-180). Springer.

Kim, C., Costello, F. J., & Lee, K. C. (2020). The unobserved heterogeneneous influence of gamification and novelty-seeking traits on consumers' repurchase intention in the omnichannel retailing. *Frontiers in psychology*, *11*.

Kim, H., & Koehler, G. (1995). Theory and practice of decision tree induction. *Omega*, *23*(6), 637-652.

Kim, J.-H. (2009). Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational Statistics & Data Analysis*, *53*(11), 3735-3745. https://doi.org/https://doi.org/10.1016/j.csda.2009.04.009

Kim, M. C., & Chen, C. (2015). A scientometric review of emerging trends and new developments in recommendation systems. *Scientometrics*, *104*(1), 239-263.

Kim, N. L., & Jin, B. E. (2020). Why buy new when one can share? Exploring collaborative consumption motivations for consumer goods. *International Journal of Consumer Studies*, *44*(2), 122-130.

King, N. (1998). Template Analysis. In G. Symon & C. Cassell (Eds.), *Qualitative Organizational Research* Sage.

Koohi, H., & Kiani, K. (2016). User based Collaborative Filtering using fuzzy C-means. *Measurement*, *91*, 134-139.

Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D., & Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks. *Machine Learning*, *95*(3), 357-380.

Kotler, P. (1974). Marketing during periods of shortage. *Journal of marketing*, *38*(3), 20-29.

Kotler, P., & Armstrong, G. (2010). *Principles of marketing*. Pearson education.

[Record #2314 is using a reference type undefined in this output style.]

Kowalski, M., Lee, Z. W., & Chan, T. K. (2021). Blockchain technology and trust relationships in trade finance. *Technological Forecasting and Social Change*, *166*, 120641.

Krisdhamara, W., Pharmasetiawan, B., & Mutijarsa, K. (2019). Improvement of Collaborative Filtering Recommendation System to Resolve Sparsity Problem using Combination of Clustering and Opinion Mining Methods. 2019 International Seminar on Application for Technology of Information and Communication (iSemantic),

Krishnamurti, T., Schwartz, D., Davis, A., Fischhoff, B., de Bruin, W. B., Lave, L., & Wang, J. (2012). Preparing for smart grid technologies: A behavioral decision research approach to understanding consumer expectations about smart meters. *Energy Policy*, *41*, 790-797.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*, 1097-1105.

Kuan, H.-H., Bock, G.-W., & Vathanophas, V. (2008). Comparing the effects of website quality on customer initial purchase and continued purchase at e-commerce websites. *Behaviour & Information Technology*, *27*(1), 3-16.

Kulkarni, N. H., Srinivasan, G., Sagar, B., & Cauvery, N. (2018). Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique. 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS),

Kumar, P., & Thakur, R. S. (2018). Recommendation system techniques and related issues: a survey. *International Journal of Information Technology*, *10*(4), 495-501.

Kumar, R. (2018). *Research methodology: A step-by-step guide for beginners*. Sage.

Lam, X. N., Vu, T., Le, T. D., & Duong, A. D. (2008). Addressing cold-start problem in recommendation systems. Proceedings of the 2nd international conference on Ubiquitous information management and communication,

Lamberton, C. P., & Rose, R. L. (2012). When is ours better than mine? A framework for understanding and altering participation in commercial sharing systems. *Journal of marketing*, *76*(4), 109-125.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

Lee, S. E., Jung, H. J., & Lee, K.-H. (2021). Motivating Collaborative Consumption in Fashion: Consumer Benefits, Perceived Risks, Service Trust, and Usage Intention of Online Fashion Rental Services. *Sustainability*, *13*(4), 1804.

Lee, S. K., Cho, Y. H., & Kim, S. H. (2010). Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations. *Information Sciences*, *180*(11), 2142-2155.

Lee, W.-J., Oh, K.-J., Lim, C.-G., & Choi, H.-J. (2014). User profile extraction from Twitter for personalized news recommendation. 16th International conference on advanced communication technology,

Lee, Z., Cheung, C., & Chan, T. K. (2021). Understanding massively multiplayer online role-playing game addiction: A hedonic management perspective. *Information systems journal*, *31*(1), 33-61.

Li, J., Jing, M., Lu, K., Zhu, L., Yang, Y., & Huang, Z. (2019). From zero-shot learning to cold-start recommendation. Proceedings of the AAAI Conference on Artificial Intelligence,

Li, J., Zhang, D., & Su, B. (2019). The impact of social awareness and lifestyles on household carbon emissions in China. *Ecological economics*, *160*, 145-155.

Li, Z., Kang, J., Yu, R., Ye, D., Deng, Q., & Zhang, Y. J. I. t. o. i. i. (2018). Consortium blockchain for secure energy trading in industrial internet of things. *14*(8), 3690-3700.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), 2674.

Lika, B., Kolomvatsos, K., & Hadjiefthymiades, S. (2014). Facing the cold start problem in recommender systems. *Expert systems with applications*, *41*(4), 2065-2073.

Lin, C.-C., Wu, H.-Y., & Chang, Y.-F. (2011). The critical factors impact on online customer satisfaction. *Procedia Computer Science*, *3*, 276-281.

Lin, C.-P., & Bhattacherjee, A. (2008). Elucidating individual intention to use interactive information technologies: The role of network externalities. *International Journal of Electronic Commerce*, *13*(1), 85-108.

Lin, P.-C., & Huang, Y.-H. (2012). The influence factors on choice behavior regarding green products based on the theory of consumption values. *Journal of Cleaner Production*, *22*(1), 11-18.

Littig, B., & Pöchhacker, F. (2014). Socio-translational collaboration in qualitative inquiry: The case of expert interviews. *Qualitative Inquiry*, *20*(9), 1085-1095.

Liu, C., Gao, X., Liu, Z., & Gao, J. (2021). When self-threat leads to the selection of emotion-enhancing options: the role of perceived transience of emotion. *European Journal of Marketing*.

Liu, F., & Lee, H. J. (2010). Use of social network information to enhance collaborative filtering performance. *Expert systems with applications*, *37*(7), 4772-4778.

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. 2008 Eighth IEEE International Conference on Data Mining,

Liu, H., Goyal, A., Walker, T., & Bhasin, A. (2014). Improving the discriminative power of inferred content information using segmented virtual profile. Proceedings of the 8th ACM Conference on Recommender systems,

Liu, J., Dolan, P., & Pedersen, E. R. (2010). Personalized news recommendation based on click behavior. Proceedings of the 15th international conference on Intelligent user interfaces,

Liu, J.-H., Zhou, T., Zhang, Z.-K., Yang, Z., Liu, C., & Li, W.-M. (2014). Promoting cold-start items in recommender systems. *PloS one*, *9*(12), e113457.

Liu, N., & Zhang, L. (2020). Green Product Design Considering Consumer’s Income and Environmental Awareness. IOP Conference Series: Earth and Environmental Science,

Liu, T., Guan, Y., & Lin, Y. (2017). Research on modulation recognition with ensemble learning. *EURASIP Journal on Wireless Communications and Networking*, *2017*(1), 1-10.

Liu, W., Deng, K., Zhang, X., Cheng, Y., Zheng, Z., Jiang, F., & Peng, J. (2020). A semi-supervised tri-catboost method for driving style recognition. *Symmetry*, *12*(3), 336.

Long, C., Wu, J., Zhang, C., Thomas, L., Cheng, M., & Jenkins, N. (2017). Peer-to-peer energy trading in a community microgrid. 2017 IEEE power & energy society general meeting,

Loper, E., & Bird, S. (2002). NLTK: the natural language toolkit. *arXiv preprint cs/0205028*.

López, V., Fernández, A., García, S., Palade, V., & Herrera, F. (2013). An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences*, *250*, 113-141.

Lu, J., Behbood, V., Hao, P., Zuo, H., Xue, S., & Zhang, G. (2015). Transfer learning using computational intelligence: a survey. *Knowledge-Based Systems*, *80*, 14-23.

Lu, J., Yao, J. E., & Yu, C.-S. (2005). Personal innovativeness, social influences and adoption of wireless Internet services via mobile technology. *The Journal of Strategic Information Systems*, *14*(3), 245-268.

Luedicke, M. K., Thompson, C. J., & Giesler, M. (2010). Consumer identity work as moral protagonism: How myth and ideology animate a brand-mediated moral conflict. *Journal of consumer research*, *36*(6), 1016-1032.

Lundberg, S. M., Erion, G. G., & Lee, S.-I. (2018). Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888*.

Luo, F., Dong, Z. Y., Liang, G., Murata, J., & Xu, Z. (2018). A distributed electricity trading system in active distribution networks based on multi-agent coalition and blockchain. *IEEE Transactions on Power Systems*, *34*(5), 4097-4108.

Luo, F., Ranzi, G., Kong, W., Liang, G., & Dong, Z. Y. (2020). Personalized residential energy usage recommendation system based on load monitoring and collaborative filtering. *IEEE transactions on industrial informatics*, *17*(2),

1253-1262.
Luo, Y., Itaya, S., Nakamura, S., & Davis, P. (2014). Autonomous cooperative energy trading between prosumers for microgrid systems. 39th Annual IEEE Conference on Local Computer Networks Workshops,
Lüth, A., Zepter, J. M., del Granado, P. C., & Egging, R. (2018). Local electricity market designs for peer-to-peer trading: The role of battery flexibility. *Applied Energy*, *229*, 1233-1243.
Ma, C., & Shen, D. i. (2021). *Guangdong's First Renewable Energy Deal Completed*. Yangcheng Everning News. Retrieved 25 Jun from https://www.china5e.com/news/news-1116939-1.html
Ma, T., Zhou, J., Tang, M., Tian, Y., Al-Dhelaan, A., Al-Rodhaan, M., & Lee, S. (2015). Social network and tag sources based augmenting collaborative recommender system. *IEICE transactions on Information and Systems*, *98*(4), 902-910.
Ma, Y., Shi, T., Zhang, W., Hao, Y., Huang, J., & Lin, Y. (2019). Comprehensive policy evaluation of NEV development in China, Japan, the United States, and Germany based on the AHP-EW model. *Journal of Cleaner Production*, *214*, 389-402.
Maamor, S., Rahman, N. L. A., & Hamed, A. B. (2016). Determinants of Islamic home financing product selection among lower income Group in Kuala Lumpur, Malaysia. *International Journal of Economics and Financial Issues*, *6*(7S).
Madani, F., Seenivasan, S., & Ma, J. (2021). Determinants of store patronage: The roles of political ideology, consumer and market characteristics. *Journal of Retailing and Consumer Services*, *63*, 102691.
Manogaran, G., Varatharajan, R., & Priyan, M. (2018). Hybrid recommendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system. *Multimedia Tools and Applications*, *77*(4), 4379-4399.
Manoharan, S. (2020). Patient Diet Recommendation System Using K Clique and Deep learning Classifiers. *Journal of Artificial Intelligence*, *2*(02), 121-130.
Marshall, A. (2009). *Principles of economics: unabridged eighth edition*. Cosimo, Inc.
Martin, C. J., & Upham, P. (2016). Grassroots social innovation and the mobilisation of values in collaborative consumption: a conceptual model. *Journal of Cleaner Production*, *134*, 204-213.
Martin, C. J., Upham, P., & Budd, L. (2015). Commercial orientation in grassroots social innovation: Insights from the sharing economy. *Ecological economics*, *118*, 240-251.
Martin, F., & Johnson, M. (2015). More Efficient Topic Modelling Through a Noun Only Approach. ALTA,
Martins, C., Oliveira, T., & Popovič, A. (2014). Understanding the Internet banking adoption: A unified theory of acceptance and use of technology and perceived risk application. *International journal of information management*, *34*(1), 1-13.
Mastrogiacomo, L., Barravecchia, F., Franceschini, F., & Marimon, F. (2021). Mining quality determinants of product-service systems from user-generated contents. *Quality Engineering*, 1-18.

Mauri, A. G., Minazzi, R., Nieto-García, M., & Viglia, G. (2018). Humanize your business. The role of personal reputation in the sharing economy. *International Journal of Hospitality Management*, *73*, 36-43.

Maxwell, J. (2008). Chapter 7: Designing a Qualitative Study. In L. Brickman & D. e. Rog (Eds.), *The SAGE handbook of applied social research methods* (Vol. 2, pp. 214-253). Sage Publication Ltd.

Mayasari, I., & Haryanto, H. C. (2018). Motivational factors of the collaborative consumption in the era of sharing economy. *Gadjah Mada International Journal of Business*, *20*(3), 331-353.

Mazaheri, E., Richard, M.-O., & Laroche, M. (2011). Online consumer behavior: Comparing Canadian and Chinese website visitors. *Journal of Business Research*, *64*(9), 958-965.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality*, *60*(2), 175-215.

McDonald, D. K. (2016). Reputation will teach the sharing economy to share. *U. Fla. JL & Pub. Pol'y*, *27*, 219.

McNeill, L., & Venter, B. (2019). Identity, self-concept and young women's engagement with collaborative, sustainable fashion consumption models. *International Journal of Consumer Studies*, *43*(4), 368-378.

Mehta, M., Rissanen, J., & Agrawal, R. (1995). MDL-Based Decision Tree Pruning. KDD,

Mengelkamp, E., Gärttner, J., Rock, K., Kessler, S., Orsini, L., & Weinhardt, C. (2018). Designing microgrid energy markets: A case study: The Brooklyn Microgrid. *Applied Energy*, *210*, 870-880.

Meuser, M., & Nagel, U. (2009). The expert interview and changes in knowledge production. In *Interviewing experts* (pp. 17-42). Springer.

Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. Proceedings of the 2011 conference on empirical methods in natural language processing,

Minami, A. L., Ramos, C., & Bortoluzzo, A. B. (2021). Sharing economy versus collaborative consumption: What drives consumers in the new forms of exchange? *Journal of Business Research*, *128*, 124-137.

Mittendorf, C. (2017). The implications of trust in the sharing economy–An empirical analysis of Uber. Proceedings of the 50th Hawaii International Conference on System Sciences,

Mittendorf, C. (2018). Collaborative consumption: the role of familiarity and trust among millennials. *Journal of Consumer Marketing*, *35*, 377-391.

Möhlmann, M. (2015). Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again. *Journal of Consumer Behaviour*, *14*(3), 193-207.

Molinaro, A. M., Simon, R., & Pfeiffer, R. M. (2005). Prediction error estimation: a comparison of resampling methods. *Bioinformatics*, *21*(15), 3301-3307. https://doi.org/10.1093/bioinformatics/bti499

Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.

Morstyn, T., Farrell, N., Darby, J., & McCulloch, D. (2018). Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants. *Nature Energy*, *3*(2), 94.

Morstyn, T., Teytelboym, A., & McCulloch, M. D. (2018). Bilateral contract networks for peer-to-peer energy trading. *IEEE Transactions on Smart Grid*, *10*(2), 2026-2035.

Motowidlo, S. J., Carter, G. W., Dunnette, M. D., Tippins, N., Werner, S., Burnett, J. R., & Vaughan, M. J. (1992). Studies of the structured behavioral interview. *Journal of applied psychology*, *77*(5), 571.

Moutinho, L. (1987). Consumer behaviour in tourism. *European Journal of Marketing*.

Myles, A. J., Feudale, R. N., Liu, Y., Woody, N. A., & Brown, S. D. (2004). An introduction to decision tree modeling. *Journal of Chemometrics: A Journal of the Chemometrics Society*, *18*(6), 275-285.

Nagarnaik, P., & Thomas, A. (2015). Survey on recommendation system methods. 2015 2nd International Conference on Electronics and Communication Systems (ICECS),

Närvänen, E., Kartastenpää, E., & Kuusela, H. (2013). Online lifestyle consumption community dynamics: A practice-based analysis. *Journal of Consumer Behaviour*, *12*(5), 358-369.

Neilson, L. A. (2010). Boycott or buycott? Understanding political consumerism. *Journal of Consumer Behaviour*, *9*(3), 214-227.

Neuman, W. L. (2014). *Basics of social research*. Pearson/Allyn and Bacon.

Neves, D., Scott, I., & Silva, C. A. (2020). Peer-to-peer energy trading potential: An assessment for the residential sector under different technology and tariff availabilities. *Energy*, *205*, 118023.

Nguyen, L. V., Hong, M.-S., Jung, J. J., & Sohn, B.-S. (2020). Cognitive similarity-based collaborative filtering recommendation system. *Applied Sciences*, *10*(12), 4183.

Nguyen, S., Peng, W., Sokolowski, P., Alahakoon, D., & Yu, X. (2018). Optimizing rooftop photovoltaic distributed generation with battery storage for peer-to-peer energy trading. *Applied Energy*, *228*, 2567-2580.

Nguyen, T. T., Harper, F. M., Terveen, L., & Konstan, J. A. (2018). User personality and user satisfaction with recommender systems. *Information Systems Frontiers*, *20*(6), 1173-1189.

Ni, S. (2021). Collaborative consumption in China: An empirical investigation of its antecedents and consequences. *Journal of Retailing and Consumer Services*, *62*, 102632.

Nikolenko, S. I., Koltcov, S., & Koltsova, O. (2017). Topic modelling for qualitative studies. *Journal of Information Science*, *43*(1), 88-102.

Nilashi, M., Bagherifard, K., Rahmani, M., & Rafe, V. (2017). A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques. *Computers & industrial engineering*, *109*, 357-368.

Ning, H., Dhelim, S., & Aung, N. (2019). PersoNet: Friend recommendation system based on big-five personality traits and hybrid filtering. *IEEE Transactions on*

Computational Social Systems, *6*(3), 394-402.
Núñez-Valdéz, E. R., Lovelle, J. M. C., Martínez, O. S., García-Díaz, V., De Pablos, P. O., & Marín, C. E. M. (2012). Implicit feedback techniques on recommender systems applied to electronic books. *Computers in Human Behavior*, *28*(4), 1186-1193.
Ocepek, U., Rugelj, J., & Bosnić, Z. (2015). Improving matrix factorization recommendations for examples in cold start. *Expert systems with applications*, *42*(19), 6784-6794.
Oh, S.-H., Lee, S. Y., & Han, C. (2020). The Effects of Social Media Use on Preventive Behaviors during Infectious Disease Outbreaks: The Mediating Role of Self-relevant Emotions and Public Risk Perception. *Health communication*, 1-10. https://doi.org/10.1080/10410236.2020.1724639
Oliveira, T., Tomar, S., & Tam, C. (2020). Evaluating collaborative consumption platforms from a consumer perspective. *Journal of Cleaner Production*, *273*, 123018.
Onwuegbuzie, A. J., & Leech, N. L. (2007). Sampling designs in qualitative research: Making the sampling process more public. *Qualitative Report*, *12*(2), 238-254.
Open Utility. (2016). *A Glimpse into the Future of Britain's Energy Economy*. https://piclo.energy/publications/piclo-trial-report.pdf
Open Utility. (2018). *Local Grid Charging-Exploring the Incentivisation of Local Energy* https://piclo.energy/publications/open-utility-local-grid-charging-white-paper.pdf
Oshio, A., Abe, S., Cutrone, P., & Gosling, S. D. (2013). Big Five content representation of the Japanese version of the Ten-Item Personality Inventory. *Psychology*, *4*(12), 924.
Paço, A., & Lavrador, T. (2017). Environmental knowledge and attitudes and behaviours towards energy consumption. *Journal of environmental management*, *197*, 384-392.
Paik, I., & Mizugai, H. (2010). Recommendation system using weighted tf-idf and naive bayes classifiers on rss contents. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *14*(6), 631-637.
Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. LREc,
Pallickara, S., & Fox, G. (2003). Naradabrokering: A distributed middleware framework and architecture for enabling durable peer-to-peer grids. ACM/IFIP/USENIX International Conference on Distributed Systems Platforms and Open Distributed Processing,
Pampel, F. (2000). *Logistic Regression: A Primer*. CA.
Park, D. H., Kim, H. K., Choi, I. Y., & Kim, J. K. (2012). A literature review and classification of recommender systems research. *Expert systems with applications*, *39*(11), 10059-10072.
Park, H. S. (2000). Relationships among attitudes and subjective norms: Testing the theory of reasoned action across cultures. *Communication Studies*, *51*(2), 162-175.

Paudel, A., Chaudhari, K., Long, C., & Gooi, H. B. (2018). Peer-to-peer energy trading in a prosumer-based community microgrid: A game-theoretic model. *IEEE Transactions on Industrial Electronics*, *66*(8), 6087-6097.

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, *7*(3), 101-134.

Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, *13*(5), 393-408.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE transactions on pattern analysis and machine intelligence*, *27*(8), 1226-1238.

Peng, M., Zhang, Q., Xing, X., Gui, T., Huang, X., Jiang, Y.-G., Ding, K., & Chen, Z. (2019). Trainable undersampling for class-imbalance learning. Proceedings of the AAAI Conference on Artificial Intelligence,

Pennacchiotti, M., & Popescu, A.-M. (2011). A machine learning approach to twitter user classification. Fifth international AAAI conference on weblogs and social media,

Perren, R., & Grauerholz, L. (2015). Collaborative consumption. *International Encyclopedia of the Social & Behavioral Sciences*, *4*(2), 139-144.

Pieroni, A., Scarpato, N., Di Nunzio, L., Fallucchi, F., Raso, M. J. I. J. o. A. S., Engineering, & Technology, I. (2018). Smarter City: Smart Energy Grid based on Blockchain Technology. *8*(1), 298-306.

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and systems magazine*, *6*(3), 21-45.

Poon, K. Y., & Huang, W.-J. (2017). Past experience, traveler personality and tripographics on intention to use Airbnb. *International Journal of Contemporary Hospitality Management*, *29*(9), 2425-2443.

Power Ledger. (2017). Where Power meets Blockchain. *Power Ledger - A New Decentralised Energy Marketplace*. https://powerledger.io/

Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.

Pretorius, A., Bierman, S., & Steel, S. J. (2016). A meta-analysis of research in random forests for classification. 2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech),

Price, B. (2002). Laddered questions and qualitative data research interviews. *Journal of advanced nursing*, *37*(3), 273-281.

Prinzie, A., & Van den Poel, D. (2008). Random forests for multiclass classification: Random multinomial logit. *Expert systems with applications*, *34*(3), 1721-1732.

Priporas, C.-V., Stylos, N., Rahimi, R., & Vedanthachari, L. N. (2017). Unraveling the diverse nature of service quality in a sharing economy. *International Journal of Contemporary Hospitality Management*.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2017). CatBoost: unbiased boosting with categorical features. *arXiv preprint arXiv:1706.09516*.

Pu, P., Chen, L., & Hu, R. (2011). A user-centric evaluation framework for recommender systems. Proceedings of the fifth ACM conference on Recommender systems,

Puschmann, T., & Alt, R. (2016). Sharing economy. *Business & Information Systems Engineering*, *58*(1), 93-99.

Quinlan, R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81-106.

Quinlan, R. (2014). *C4. 5: programs for machine learning*. Elsevier.

Quintal, V. A., Lee, J. A., & Soutar, G. N. (2010). Risk, uncertainty and the theory of planned behavior: A tourism example. *Tourism Management*, *31*(6), 797-805.

Rahm, E., & Do, H. H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.*, *23*(4), 3-13.

Rai, V., & Robinson, S. A. (2015). Agent-based modeling of energy technology adoption: Empirical integration of social, behavioral, economic, and environmental factors. *Environmental Modelling & Software*, *70*, 163-177.

Rajak, R. K., Pawar, A., Pendke, M., Shinde, P., Rathod, S., & Devare, A. (2017). Crop recommendation system to maximize crop yield using machine learning technique. *International Research Journal of Engineering and Technology*, *4*(12), 950-953.

Ramakrishnan, G., Saicharan, V., Chandrasekaran, K., Rathnamma, M., & Ramana, V. V. (2020). Collaborative filtering for book recommendation system. In *Soft Computing for Problem Solving* (pp. 325-338). Springer.

Ramírez-Esparza, N., Gosling, S. D., Benet-Martínez, V., Potter, J. P., & Pennebaker, J. W. (2006). Do bilinguals have two personalities? A special case of cultural frame switching. *Journal of research in personality*, *40*(2), 99-120.

Rashid, A. M., Albert, I., Cosley, D., Lam, S. K., McNee, S. M., Konstan, J. A., & Riedl, J. (2002). Getting to know you: learning new user preferences in recommender systems. Proceedings of the 7th international conference on Intelligent user interfaces,

Rathore, S. S., & Kumar, S. (2017). A decision tree logic based recommendation system to select software fault prediction techniques. *Computing*, *99*(3), 255-285.

Ray, A., Bala, P. K., & Dwivedi, Y. K. (2021). Exploring values affecting e-Learning adoption from the user-generated-content: A consumption-value-theory perspective. *Journal of Strategic Marketing*, *29*(5), 430-452.

Redman, T. C. (1998). The impact of poor data quality on the typical enterprise. *Communications of the ACM*, *41*(2), 79-82.

Řehůřek, R., & Sojka, P. (2011). Gensim—statistical semantics in python. *Retrieved from genism. org*.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of netnews. Proceedings of the 1994 ACM conference on Computer supported cooperative work,

Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3), 56-58.

Retamal, M. (2019). Collaborative consumption practices in Southeast Asian cities: Prospects for growth and sustainability. *Journal of Cleaner Production*, *222*,

143-152.
Reyes-Menendez, A., Saura, J. R., & Martinez-Navalon, J. G. (2019). The impact of e-WOM on hotels management reputation: exploring tripadvisor review credibility with the ELM model. *Ieee Access*, *7*, 68868-68877.
Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In *Recommender systems handbook* (pp. 1-35). Springer.
Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: introduction and challenges. In *Recommender systems handbook* (pp. 1-34). Springer.
Rinaldi, A., Schweiker, M., & Iannone, F. (2018). On uses of energy in buildings: Extracting influencing factors of occupant behaviour by means of a questionnaire survey. *Energy and buildings*, *168*, 298-308.
Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93-104. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2011.11.002
Rogers, E. M. (2002). Diffusion of preventive innovations. *Addictive behaviors*, *27*(6), 989-993.
Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, *33*(1-2), 1-39.
Roos, D., & Hahn, R. (2017). Does shared consumption affect consumers' values, attitudes, and norms? A panel study. *Journal of Business Research*, *77*, 113-123.
Rosen, D., Lafontaine, P. R., & Hendrickson, B. (2011). CouchSurfing: Belonging and trust in a globally cooperative online social network. *New Media & Society*, *13*(6), 981-998.
Roy, S. K., Balaji, M., Soutar, G., Lassar, W. M., & Roy, R. (2018). Customer engagement behavior in individualistic and collectivistic markets. *Journal of Business Research*, *86*, 281-290.
Ryan, J., & Casidy, R. (2018). The role of brand reputation in organic food consumption: A behavioral reasoning perspective. *Journal of Retailing and Consumer Services*, *41*, 239-247.
Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, *21*(3), 660-674.
Safoury, L., & Salah, A. (2013). Exploiting user demographic attributes for solving cold-start problem in recommender system. *Lecture Notes on Software Engineering*, *1*(3), 303-307.
Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(4), e1249.
Saha, J., Chowdhury, C., & Biswas, S. (2020). Review of machine learning and deep learning based recommender systems for health informatics. In *Deep Learning Techniques for Biomedical and Health Informatics* (pp. 101-126). Springer.
Sahebi, S., & Cohen, W. W. (2011). Community-based recommendations: a solution to the cold start problem. Workshop on recommender systems and the social web,

RSWEB,
Samukange, W. T., Gardarsdottir, H., Leufkens, H. G., & Mantel-Teeuwisse, A. K. (2020). Selection of Blood, Blood Components, and Blood Products as Essential Medicines in 105 Low-and Middle-Income Countries. *Transfusion medicine reviews*, *34*(2), 94-100.
Sardianou, E., & Genoudi, P. (2013). Which factors affect the willingness of consumers to adopt renewable energies? *Renewable energy*, *57*, 1-4.
Saveski, M., & Mantrach, A. (2014). Item cold-start recommendations: learning local collective embeddings. Proceedings of the 8th ACM Conference on Recommender systems,
Sawrikar, P., & Katz, I. B. (2014). Recommendations for improving cultural competency when working with ethnic minority families in child protection systems in Australia. *Child and Adolescent Social Work Journal*, *31*(5), 393-417.
Scapolo, F., & Miles, I. (2006). Eliciting experts' knowledge: A comparison of two methods. *Technological Forecasting and Social Change*, *73*(6), 679-704.
Scaraboto, D. (2015). Selling, sharing, and everything in between: The hybrid economies of collaborative networks. *Journal of consumer research*, *42*(1), 152-176.
Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. *Data mining and knowledge discovery*, *5*(1-2), 115-153.
Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, *5*(2), 197-227.
Schedl, M., Zamani, H., Chen, C.-W., Deldjoo, Y., & Elahi, M. (2018). Current challenges and visions in music recommender systems research. *International Journal of Multimedia Information Retrieval*, *7*(2), 95-116.
Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. (2002). Methods and metrics for cold-start recommendations. Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval,
Schmitt-Beck, R. (2015). Bandwagon effect. *The international encyclopedia of political communication*, 1-5.
Schultze, U., & Avital, M. (2011). Designing interviews to generate rich data for information systems research. *Information and organization*, *21*(1), 1-16.
Seddon, P. B. (1997). A respecification and extension of the DeLone and McLean model of IS success. *Information systems research*, *8*(3), 240-253.
Sedera, D., Lokuge, S., Atapattu, M., & Gretzel, U. (2017). Likes—the key to my happiness: The moderating effect of social influence on travel experience. *Information & management*, *54*(6), 825-836.
Sen, A. (2009). *The Idea of Justice*. (London: Allen Lane)
Seyfang, G., Hielscher, S., Hargreaves, T., Martiskainen, M., & Smith, A. (2014). A grassroots sustainable energy niche? Reflections on community energy in the UK. *Environmental Innovation and Societal Transitions*, *13*, 21-44.
Shabbir, M. S., Bait Ali Sulaiman, M. A., Hasan Al-Kumaim, N., Mahmood, A., &

Abbas, M. (2020). Green Marketing Approaches and Their Impact on Consumer Behavior towards the Environment—A Study from the UAE. *Sustainability*, *12*(21), 8977.

Shahabi, C., Banaei-Kashani, F., Chen, Y.-S., & McLeod, D. (2001). Yoda: An accurate and scalable web-based recommendation system. International Conference on Cooperative Information Systems,

Shannon, E. (2001). A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review*, *5*(1), 3-55.

Shao, G. (2009). Understanding the appeal of user-generated media: a uses and gratification perspective. *Internet Research*.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, *2*(28), 307-317.

Sheth, J. N. (1973). *A field study of attitude structure and attitude-behavior relationship*. College of Commerce and Business Administration, University of Illinois at ….

Sheth, J. N., Newman, B. I., & Gross, B. L. (1991). Why we buy what we buy: A theory of consumption values. *Journal of Business Research*, *22*(2), 159-170.

Singelis, T. M., Triandis, H. C., Bhawuk, D. P., & Gelfand, M. J. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-cultural research*, *29*(3), 240-275.

Sinha, I., & DeSarbo, W. S. (1998). An integrated approach toward the spatial modeling of perceived customer value. *Journal of Marketing Research*, *35*(2), 236-249.

Smith, A., Hargreaves, T., Hielscher, S., Martiskainen, M., & Seyfang, G. (2016). Making the most of community energies: Three perspectives on grassroots innovation. *Environment and Planning A*, *48*(2), 407-432.

Smith, A. N., Fischer, E., & Yongjian, C. (2012). How does brand-related user-generated content differ across YouTube, Facebook, and Twitter? *Journal of interactive marketing*, *26*(2), 102-113.

Sobhanam, H., & Mariappan, A. (2013a). Addressing cold start problem in recommender systems using association rules and clustering technique. 2013 International Conference on Computer Communication and Informatics,

Sobhanam, H., & Mariappan, A. (2013b). A hybrid approach to solve cold start problem in recommender systems using association rules and clustering technique. *Int. J. Comput. Appl*, *74*(4), 17-23.

Sommerfeld, J., Buys, L., Mengersen, K., & Vine, D. (2017). Influence of demographic variables on uptake of domestic solar photovoltaic technology. *Renewable and Sustainable Energy Reviews*, *67*, 315-323.

Sonnenbatterie. (2020). *What Is the SonnenCommunity?* https://sonnengroup.com/sonnencommunity/

Sordi, J. D., Perin, M. G., Petrini, M. d. C., & Sampaio, C. H. (2018). Construal level and collaborative consumption: An exploratory approach. *International Journal of Consumer Studies*, *42*(2), 264-273.

Souri, A., Hosseinpour, S., & Rahmani, A. M. (2018). Personality classification based on profiles of social networks' users and the five-factor model of personality. *Human-centric Computing and Information Sciences*, *8*(1), 1-15.

Sousa, T., Soares, T., Pinson, P., Moret, F., Baroche, T., & Sorin, E. (2019). Peer-to-peer and community-based markets: A comprehensive review. *Renewable and Sustainable Energy Reviews*, *104*, 367-378.

Sovacool, B. K., & Valentine, S. V. (2010). The socio-political economy of nuclear energy in China and India. *Energy*, *35*(9), 3803-3813.

Sovacool, B. K., Valentine, S. V., Bambawale, M. J., Brown, M. A., de Fátima Cardoso, T., Nurbek, S., Suleimenova, G., Li, J., Xu, Y., & Jain, A. (2012). Exploring propositions about perceptions of energy security: An international survey. *Environmental Science & Policy*, *16*, 44-64.

Steelman, Z. R., Hammer, B. I., & Limayem, M. (2014). Data collection in the digital age. *MIS quarterly*, *38*(2), 355-378.

Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, *62*(1), 77-89. https://doi.org/https://doi.org/10.1016/S0034-4257(97)00083-7

Stigler, G. J. (1950). The development of utility theory. I. *Journal of political Economy*, *58*(4), 307-327.

Stone, T., & Choi, S.-K. (2013). Extracting consumer preference from user-generated content sources using classification. International Design Engineering Technical Conferences and Computers and Information in Engineering Conference,

Sugandini, D., Feriyanto, N., Muafi, M., Hadioetomo, H., & Hapsoro Darpito, S. (2018). The influence of novelty seeking behavior and autonomy toward new product trial in the context of e-business learning: The role of mediation of perceived behavioral control. *Journal of Entrepreneurship Education*, *21*(3), 1-9.

Sun, M., Li, F., Lee, J., Zhou, K., Lebanon, G., & Zha, H. (2013). Learning multiple-question decision trees for cold-start recommendation. Proceedings of the sixth ACM international conference on Web search and data mining,

Sun, Y., Wong, A. K., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International journal of pattern recognition and artificial intelligence*, *23*(04), 687-719.

[Record #579 is using a reference type undefined in this output style.]

Suresh, G., Kumar, A. S., Lekashri, S., & Manikandan, R. (2021). Efficient crop yield recommendation system using machine learning for digital farming. *International Journal of Modern Agriculture*, *10*(1), 906-914.

Susarla, A., Oh, J.-H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information systems research*, *23*(1), 23-41.

Swanson, E. B. (1974). Management information systems: appreciation and involvement. *Management science*, *21*(2), 178-188.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*(4857), 1285. https://doi.org/10.1126/science.3287615

Syed, S., & Spruit, M. (2017). Full-text or abstract? examining topic coherence scores using latent dirichlet allocation. 2017 IEEE International conference on data science and advanced analytics (DSAA), Taleb, I., Dssouli, R., & Serhani, M. A. (2015). Big data pre-processing: A quality framework. 2015 IEEE international congress on big data,

Talia, D., & Trunfio, P. (2003). Toward a synergy between p2p and grids. *IEEE Internet Computing*, *7*(4), 96-95.

Talwar, S., Dhir, A., Kaur, P., & Mäntymäki, M. (2020). Why do people purchase from online travel agencies (OTAs)? A consumption values perspective. *International Journal of Hospitality Management*, *88*, 102534.

Tan, F. B., & Hunter, M. G. (2002). The repertory grid technique: A method for the study of cognition in information systems. *MIS quarterly*, 39-57.

Tang, L., Jiang, Y., Li, L., & Li, T. (2014). Ensemble contextual bandits for personalized recommendation. Proceedings of the 8th ACM Conference on Recommender Systems,

[Record #422 is using a reference type undefined in this output style.]

Ter Huurne, M., Ronteltap, A., Guo, C., Corten, R., & Buskens, V. (2018). Reputation effects in socially driven sharing economy transactions. *Sustainability*, *10*(8), 2674.

Tey, F. J., Wu, T.-Y., Lin, C.-L., & Chen, J.-L. (2021). Accuracy improvements for cold-start recommendation problem using indirect relations in social networks. *Journal of Big Data*, *8*(1), 1-18.

Tharwat, A. (2020). Classification assessment methods. *Applied Computing and Informatics*, *17*(1), 168-192.

Thiengburanathum, P., Cang, S., & Yu, H. (2015). A decision tree based recommendation system for tourists. 2015 21st International Conference on Automation and Computing (ICAC),

ThØgersen, J. (1995). Understanding of consumer behaviour as a prerequisite for environmental protection. *Journal of consumer policy*, *18*(4), 345-385.

Thøgersen, J., & Ölander, F. (2003). Spillover of environment-friendly consumer behaviour. *Journal of environmental psychology*, *23*(3), 225-236.

Thomas, L., Zhou, Y., Long, C., Wu, J., & Jenkins, N. (2019). A general form of smart contract for decentralized energy systems management. *Nature Energy*, *4*(2), 140-149.

Timoshenko, A., & Hauser, J. R. (2019). Identifying customer needs from user-generated content. *Marketing Science*, *38*(1), 1-20.

Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research*, *51*(4), 463-479.

Tkalcic, M., & Chen, L. (2015). Personality and recommender systems. In *Recommender systems handbook* (pp. 715-739). Springer.

Toni, M., Renzi, M. F., & Mattia, G. (2018). Understanding the link between collaborative economy and sustainable behaviour: An empirical investigation. *Journal of Cleaner Production*, *172*, 4467-4477.

Triandis, H. C., & Suh, E. M. (2002). Cultural influences on personality. *Annual review of psychology*, *53*(1), 133-160.

Tripoliti, E. E., Fotiadis, D. I., & Manis, G. (2013). Modifications of the construction and voting mechanisms of the random forests algorithm. *Data & Knowledge Engineering*, *87*, 41-65.

Trope, Y., Liberman, N., & Wakslak, C. (2007). Construal levels and psychological distance: Effects on representation, prediction, evaluation, and behavior. *Journal of consumer psychology*, *17*(2), 83-95.

Trunfio, P., Talia, D., Papadakis, H., Fragopoulou, P., Mordacchini, M., Pennanen, M., Popov, K., Vlassov, V., & Haridi, S. (2007). Peer-to-Peer resource discovery in Grids: Models and systems. *Future Generation Computer Systems*, *23*(7), 864-878.

Tsai, C.-F., & Hung, C. (2012). Cluster ensembles in collaborative filtering recommendation. *Applied Soft Computing*, *12*(4), 1417-1425.

Tunçel, N., & Özkan Tektaş, Ö. (2020). Intrinsic motivators of collaborative consumption: A study of accommodation rental services. *International Journal of Consumer Studies*, *44*(6), 616-628.

Tushar, W., Saha, T. K., Yuen, C., Azim, M. I., Morstyn, T., Poor, H. V., Niyato, D., & Bean, R. (2020). A coalition formation game framework for peer-to-peer energy trading. *Applied Energy*, *261*, 114436.

Tussyadiah, I. P. (2016). Factors of satisfaction and intention to use peer-to-peer accommodation. *International Journal of Hospitality Management*, *55*, 70-80.

Tussyadiah, I. P., & Pesonen, J. (2018). Drivers and barriers of peer-to-peer accommodation stay–an exploratory study with American and Finnish travellers. *Current issues in Tourism*, *21*(6), 703-720.

Ullah, M. H., & Park, J.-D. (2021). Peer-to-peer energy trading in transactive markets considering physical network constraints. *IEEE Transactions on Smart Grid*, *12*(4), 3390-3403.

Van den Bulck, H., Puppis, M., Donders, K., & Van Audenhove, L. (2019). *The Palgrave Handbook of Methods for Media Policy Research*. Springer.

Vanderbron. (2017). *White Paper of Vanderbron*. https://vandebron.nl

Varshneya, G., Pandey, S. K., & Das, G. (2017). Impact of social influence and green consumption values on purchase intention of organic clothing: a study on collectivist developing economy. *Global Business Review*, *18*(2), 478-492.

Veblen, T., & Galbraith, J. K. (1973). *The theory of the leisure class*. Houghton Mifflin Boston.

Venkatesan, M., & Thangadurai, K. (2017). History and Overview of the Recommender Systems. In *Collaborative Filtering Using Data Mining and Analysis* (pp. 74-99). IGI Global.

Victor, P., Cornelis, C., Teredesai, A. M., & De Cock, M. (2008). Whom should I trust? The impact of key figures on cold start recommendations. Proceedings of the 2008 ACM symposium on Applied computing,

Viktoratos, I., Tsadiras, A., & Bassiliades, N. (2018). Combining community-based knowledge with association rule mining to alleviate the cold start problem in context-aware recommender systems. *Expert systems with applications*, *101*, 78-90.

Vincent, D. R., Deepa, N., Elavarasan, D., Srinivasan, K., Chauhdary, S. H., & Iwendi,

C. (2019). Sensors driven AI-based agriculture recommendation model for assessing land suitability. *Sensors*, *19*(17), 3667.

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.-A., & Bottou, L. (2010). Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine Learning research*, *11*(12).

Vroom, V. H. (1964). Work and motivation.

Waheed, T., Bonnell, R., Prasher, S. O., & Paulet, E. (2006). Measuring performance in precision agriculture: CART—A decision tree approach. *Agricultural water management*, *84*(1-2), 173-185.

Walker, G., & Devine-Wright, P. (2008). Community renewable energy: What should it mean? *Energy Policy*, *36*(2), 497-500.

Wan, L., Hong, Y., Huang, Z., Peng, X., & Li, R. (2018). A hybrid ensemble learning method for tourist route recommendations based on geo-tagged social networks. *International Journal of Geographical Information Science*, *32*(11), 2225-2246.

Wang, D., Liang, Y., Xu, D., Feng, X., & Guan, R. (2018). A content-based recommender system for computer science publications. *Knowledge-Based Systems*, *157*, 1-9.

Wang, H., Wang, N., & Yeung, D.-Y. (2015). Collaborative deep learning for recommender systems. Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining,

Wang, J., & Wu, L. (2016). The impact of emotions on the intention of sustainable consumption choices: evidence from a big city in an emerging country. *Journal of Cleaner Production*, *126*, 325-336.

Wang, N., Zhou, X., Lu, X., Guan, Z., Wu, L., Du, X., & Guizani, M. (2019). When energy trading meets blockchain in electrical power system: The state of the art. *Applied Sciences*, *9*(8), 1561.

Wang, Q., & Sun, X. (2017). Crude oil price: Demand, supply, economic activity, economic policy uncertainty and wars–From the perspective of structural equation modelling (SEM). *Energy*, *133*, 483-490.

Wang, S. I., & Manning, C. D. (2012). Baselines and Bigrams: Simple, Good Sentiment and Topic Classification. ACL,

Wang, X., Peng, Z., Wang, S., Philip, S. Y., Fu, W., Xu, X., & Hong, X. (2019). CDLFM: cross-domain recommendation for cold-start users via latent feature mapping. *Knowledge and Information Systems*, 1-28.

Wang, Y., Singgih, M., Wang, J., & Rit, M. (2019). Making sense of blockchain technology: How will it transform supply chains? *International Journal of Production Economics*, *211*, 221-236.

Wang, Y., Wu, H., Lei, X., Shen, J., & Feng, Z. (2020). The Influence of Doctors' Online Reputation on the Sharing of Outpatient Experiences: Empirical Study. *Journal of medical Internet research*, *22*(12), e16691.

Wang, Y.-S., & Liao, Y.-W. (2008). Assessing eGovernment systems success: A validation of the DeLone and McLean model of information systems success. *Government information quarterly*, *25*(4), 717-733.

Wang, Y.-S., & Tang, T.-I. (2003). Assessing customer perceptions of website service quality in digital marketing environments. *Journal of Organizational and End User Computing (JOEUC)*, *15*(3), 14-31.

Warriner, G. K. (1981). Electricity consumption by the elderly: policy implications. *Journal of consumer research*, *8*(3), 258-264.

Weber, L. (1997). Some reflections on barriers to the efficient use of energy. *Energy Policy*, *25*(10), 833-835.

Wei, J., He, J., Chen, K., Zhou, Y., & Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert systems with applications*, *69*, 29-39.

Wei, X., Lo, C. K., Jung, S., & Choi, T.-M. (2021). From co-consumption to co-production: A systematic review and research synstudy of collaborative consumption practices. *Journal of Business Research*, *129*, 282-294.

Wen, X. (2021). Using deep learning approach and IoT architecture to build the intelligent music recommendation system. *Soft Computing*, *25*(4), 3087-3096.

Wiggins, J. S. (1996). *The five-factor model of personality: Theoretical perspectives*. Guilford Press.

Wilkins, D. J., Chitchyan, R., & Levine, M. (2020). Peer-to-peer energy markets: Understanding the values of collective and community trading. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,

Witten, I. H., Paynter, G. W., Frank, E., Gutwin, C., & Nevill-Manning, C. G. (2005). Kea: Practical automated keyphrase extraction. In *Design and Usability of Digital Libraries: Case Studies in the Asia Pacific* (pp. 129-152). IGI global.

Wittkowski, K., Moeller, S., & Wirtz, J. (2013). Firms' intentions to use nonownership services. *Journal of Service Research*, *16*(2), 171-185.

Woodall, T. (2003). Conceptualising 'value for the customer': an attributional, structural and dispositional analysis. *Academy of marketing science review*, *12*(1), 1-42.

Woźniak, M., Graña, M., & Corchado, E. (2014). A survey of multiple classifier systems as hybrid systems. *Information Fusion*, *16*, 3-17.

Wu, Q., Burges, C. J., Svore, K. M., & Gao, J. (2010). Adapting boosting for information retrieval measures. *Information Retrieval*, *13*(3), 254-270.

Yan, D., O'Brien, W., Hong, T., Feng, X., Gunay, H. B., Tahmasebi, F., & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and buildings*, *107*, 264-278.

Yang, J., Paudel, A., & Gooi, H. (2019). Blockchain framework for peer-to-peer energy trading with credit rating. 2019 IEEE Power & Energy Society General Meeting (PESGM),

Yao, B., Qiu, R. T., Fan, D. X., Liu, A., & Buhalis, D. (2019). Standing out from the crowd–an exploration of signal attributes of Airbnb listings. *International Journal of Contemporary Hospitality Management*.

Yera, R., & Martinez, L. (2017). Fuzzy tools in recommender systems: A survey. *International Journal of Computational Intelligence Systems*, *10*(1), 776-803.

Yu, Z. J., Haghighat, F., Fung, B. C., Morofsky, E., & Yoshino, H. (2011). A methodology for identifying and improving occupant behavior in residential buildings. *Energy*, *36*(11), 6596-6608.

Zajonc, R. (1968). ``Attitudinal effects of mere exposure''Journal of Personality and Social Psychology 9 1^ 27 Conditions of use. *This article may be downloaded from the E&P website for personal research by members of subscribing organisations. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher*.

Zarifis, A., Cheng, X., & Kroenung, J. (2019). Collaborative consumption for low and high trust requiring business models: from fare sharing to supporting the elderly and people with disability. *International Journal of Electronic Business*, *15*(1), 1-20.

Zeba, F., & Ganguli, S. (2019). Novelty seeking as driving factor for fashion apparel innovators: A qualitative investigation. *Journal of Global Fashion Marketing*, *10*(3), 246-266.

Zhang, C., Wu, J., Long, C., & Cheng, M. (2017). Review of existing peer-to-peer energy trading projects. *Energy Procedia*, *105*, 2563-2568.

Zhang, C., Wu, J., Zhou, Y., Cheng, M., & Long, C. (2018). Peer-to-Peer energy trading in a Microgrid. *Applied Energy*, *220*, 1-12.

Zhang, K., Chen, Y., & Li, C. (2019). Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing. *Tourism Management*, *75*, 595-608.

Zhang, K., Chen, Y., & Lin, Z. (2020). Mapping destination images and behavioral patterns from user-generated photos: a computer vision approach. *Asia Pacific Journal of Tourism Research*, *25*(11), 1199-1214.

Zhang, M., Tang, J., Zhang, X., & Xue, X. (2014). Addressing cold start in recommender systems: A semi-supervised co-training algorithm. Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval,

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys (CSUR)*, *52*(1), 1-38.

Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies*, *58*, 308-324.

Zhang, Y., Yang, C., & Niu, Z. (2014). A research of job recommendation system based on collaborative filtering. 2014 Seventh International Symposium on Computational Intelligence and Design,

Zhang, Z., Liu, Y., Xu, G., & Chen, H. (2016). A weighted adaptation method on learning user preference profile. *Knowledge-Based Systems*, *112*, 114-126.

Zhang, Z.-K., Liu, C., Zhang, Y.-C., & Zhou, T. (2010). Solving the cold-start problem in recommender systems with social tags. *EPL (Europhysics Letters)*, *92*(2), 28002.

Zhao, M., & Xie, J. (2011). Effects of social and temporal distance on consumers' responses to peer recommendations. *Journal of Marketing Research*, *48*(3), 486-496.

Zhao, S., Song, Q., & Wang, C. (2019). Characterizing the energy-saving behaviors, attitudes and awareness of university students in Macau. *Sustainability*, *11*(22), 6341.

Zhao, W. X., Li, S., He, Y., Chang, E. Y., Wen, J.-R., & Li, X. (2015). Connecting social media to e-commerce: Cold-start product recommendation using microblogging information. *IEEE transactions on knowledge and data engineering*, *28*(5), 1147-1159.

Zhao, Z.-D., & Shang, M.-S. (2010). User-based collaborative-filtering recommendation algorithms on hadoop. 2010 third international conference on knowledge discovery and data mining,

Zhong, W., Yu, R., Xie, S., Zhang, Y., & Tsang, D. H. (2016). Software defined networking for flexible and green energy internet. *IEEE Communications Magazine*, *54*(12), 68-75.

Zhou, K., Yang, S.-H., & Zha, H. (2011). Functional matrix factorizations for cold-start recommendation. Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval,

Zhou, Y., Wu, J., & Long, C. (2018). Evaluation of peer-to-peer energy sharing mechanisms based on a multiagent simulation framework. *Applied Energy*, *222*, 993-1022.

Zhou, Z.-H. (2009). Ensemble learning. *Encyclopedia of biometrics*, *1*, 270-273.

Zhuang, F., Zhou, Y., Zhang, F., Ao, X., Xie, X., & He, Q. (2017). Sequential transfer learning: Cross-domain novelty seeking trait mining for recommendation. Proceedings of the 26th international conference on world wide web companion,

Zihayat, M., Ayanso, A., Zhao, X., Davoudi, H., & An, A. (2019). A utility-based news recommendation system. *Decision support systems*, *117*, 14-27.

APPENDIXES

Appendix A: Interview Schedule

谢谢你接受我的邀请。现在越来越多的能源消费者现在变成了生产者。由于包括光伏电池和风力涡轮机等分布式能源资源的创新，用户在使用能源的同时生产能源。这就是创造点对点能源交易平台最基础的动力。我正在进行一项研究，了解客户在 P2P 能源交易的产品上的表现。我有几个问题想问你，感谢你的回答。

你能介绍一下基于自己领域的 P2P 能源交易项目吗？

i. 你能描述一下目前 p2p 能源的设置/价格机制/市场机制/信息系统/监管是如何运作的吗？
ii. 你认为目前的能源设置/价格机制/市场机制/信息系统/监管方面的主要挑战是什么？
iii. 在未来的 P2P 能源交易中，是否存在任何挑战和机遇？

PROFILE	DATA
M1	关于现在中国的能源交易电网的搭建，主要是建立在农业用电，主网相连。用户可以自给自足，也可以反向卖给主网。 关于能源交易本身而言，不同职能部门专注点肯定不同，对于电力运营部门，关注的是电的就地消纳问题，以及怎么减少运输过程中的电力损耗。还有，我们需要考虑新能源本身就是不稳定的，有太阳和没有太阳差别很大，有风和没风，差距也很大。因此在他们并入电网的时候对电网本身造成的波动性也很大。 因此很多顾客在初次使用的时候，就会问我们，是否并网。也就是，是否有加入主网？因为他们很害怕自己的发的电不够用的，怕自己要承担风险。不连入主网会让他们感觉电力供给不稳定.还有并入主网，也是国家提倡和要求的，农户可以赚取额外的电费补贴。而且很多农户用电很节省，基本可以实现自给自足。由于中国地大物博，东西光电互补差距很大，在宁夏这些地方，发的电根本用不完。 所以才有了西电东输，西气东送。 对于市场而言，现在中国的能源交易主要有三种。一种是农光互补。也就是政府对使用太阳能发电的农户有一定的经济补贴。另外一种是区块链分布式账簿技术和人工智能为支撑、基于太阳能和电池储能的电力交易平台。比如说，荷兰的公司 distro 在做这方面的研究。在这种情况下，消费者更能清楚的知道能源的流向。买卖双方以此了解市场偏好、供需情况和能源价格动态信息，进而更自由、更高效地促成电力买卖的达成。现在这些平台基本都在向区块链为基础的平台转型。每个市场参与者都配备了一个充当“交易代理工具”的人工智能仪，从根本上去除中间商，节约时间和成本，促成交易。 此外区块链可以有效降低信任成本，提高交易效率，安全可靠；而分布式能源的去中心化、开放共享、自主性和安全性，符合能源的发展趋势。区块链在能源领域的应用，将有效支持多种类型能源系统的开放互联和多种用户的广泛深入参与。通过共同维护可信的分布式账本，可以实现未来能源交易中能量流、信息流和价值链的有效连接。

对于用户来说，他们的使用的热情度还是很高的，因为国家有补贴，政策友好，他们也节省了开支。 再就是大型发电农场，比如说风力电厂，一般都建立在郊区，有一大块地，收集了电以后，卖给一些企业，实现双向受益。

在国外，现在很多企业都加入了区块链平台。 通过区块链和邻居， 社区进行交易。咱们首先不分析区块链好不好，现在什么产品都讲究上链。 说是可以实现匿名，更加安全。

比如说澳大利亚的一个公司，在市区推出以区块链为基础的家庭剩余能源交易的系统。当电力生产商进入并向配电网络出售绿色电力时，配电网络运营商会对此进行确认，确认成功后，智能合约会生成相应的虚拟货币并提供给绿色电力生产商作为奖励。整个交易很简单，通过使用虚拟货币提高了信任度。

而且有些公司还推出了能源币，其实就是建立在以太坊，或者比特币基础上的代币。顾客其实对这种交易模式并不是很认可。因为能源币浮动大吗，和传统的美元结算，他们貌似并不受益。 今天比特币 4 万美元，明天可能长了 1000 美元。所以这种交易对于能源本身的交易并不对等。

再者，加密货币不受政府和中央银行的管制。这一点真的很麻烦。在中国现有的政治意识形态下，货币肯定不可能失去中央政府的监管。中国人很聪明的，如果放开监管，相信到时非法交易会层出不穷。当然，我觉得西方社会也不会将这种加密货币合法化。虽说西方国家的意识形态和中国有所不同，但是最终目的都是一样的。西方国家可能在对待加密货币的认识上有所不同，他们可能认为这是一种商品。可以用开放的态度来看这种商品。但是在中国，这种加密货币对我们的货币系统造成了威胁，它威胁了国家银行的职能地位。所以在中国，加密货币不可能作为法定货币。而能源币也不可能进行自由交易。除非亿人民币直接进行结算。或是中央银行推出的电子货币直接结算。

现在国外的一些平台也推出一些优惠套餐，比如说安装太阳光板，可以免费赠送相应数量的能源币，或者在下次购买时，可以拿到一定的折扣。其实这些对消费者而言诱惑并不大，消费者考虑的首先是我买谁的，我卖给谁，靠谱吗？我不用了，能退款吗？能源不同于其他的产品，本身他就是一个硬性的需求。消费者才不管你是什么电，反正到最后都是一样的电，都可以炒菜做饭，看电视。而且这种折扣和安装费相比，能不能在几年内，得到有效回报。太阳板光板是有年限限制的，一般 20 年就需要更换。那么后续的问题是，你又要投入新的成本来更新光板。但关键是在这 20 年里，拟投入的成本有没有回本。如果你住的地方阳光足够充沛，一年四季都有充足的阳光，那可以考虑。所以这个项目在澳大利亚要长效一些。如果是英国一年有 200 天下雨，可能就不适合。

再有，顾客在买东西的时候，会考虑这个品牌的信誉，形象等各种问题，才会慢慢培养起忠诚度。能源是一个道理。很多公司都在做这个可持续能源的售卖。那么有的顾客就会特别重视这个公司的信誉，要看是不是大公司，值不值得信赖。

还有就是，消费者关注的更多的是普及程度，如果我的邻居他们都安装了，我是不是要安装，这样在一个社区里，融入性是不是更好。换句话说，如果小区的的入户普及率高的话，消费者更能找到归属感。再有，你说你最好的朋友告诉你，他们家安装了这个产品，可以和别人进行交易了。你会不会也想安装。如果他再和你说，有多么优惠，你会不会更加动心。

现在有什么挑战，主要是对于建立这个点对点交易的平台方面，这种端对端能源交易在协商、执行和清算等各个阶段均涉及大量的通信和信息。对集中式和分布式端对端能源交易市场而言，市场协调者与各个用户间需要大量的双边通信；对分散式市场而

	言，各个用户间也需要大量的通信。未来研究方面，通信负荷和投资需求可以作为评价端对端能源交易市场机制和交易平台的一个重要维度。非理想通信条件，如通信延迟、畸变和失败对端对端能源交易的影响有待进一步评估。此外，面向端对端能源交易的通信网络优化配置和运行也有待进一步研究。 还有，其实在现实中很多端对端的价格甚至高于市场价格。有研究表明，人们对自足性的偏好将降低端对端能源交易的可能性，因为他们认为这使得人们在给自己的本地发电定价时会倾向于给出高于其实际价值的价格。换句话说就是自己的用电量满足了，实现了自己对电的控制权，用户对产品的实际价值预期并没有那么高。 所以我们所说的用户最看重的经济利益也许并不成立。还有，如果在一个社区里，你的邻居都不使用，相信你自己使用的可能性也不大。因为你还要承担，电力不稳定的风险。对于能源政策来说，这一点我相信中国是做的很好的。咱们国家很支持。从十三五到现在， 国家对农业和渔业，光能互补发电这一块咱们做得很好。 具体到国外，咱们和国外的形式不一样。国外电力私有化，咱们是国有化。咱们是国家统一调度。 所以在未来，如果邻居之间的电力交易，应该有很长的路要走。而且你看看中国现在建这么多高楼，也没地方放太阳能板啊。放上几块小太阳能板，也不够整栋的居民使用。我们用的更多的是太阳能热水器. 至于对于未来发展，我特别想提一下区块链的问题，现在区块链在中国很火，到处都讲区块链，都讲上链。具体区块链怎么工作？一般在信息发布和汇总阶段，具有合格的电力资质用户和分布式能源发电商通过网站或应用程序注册并提出购电和售电请求，用户在应用层达成交易意向，即买卖双方的匹配在此阶段完成，无需在区块链中进行匹配。然后向区块链传输信息。在应用框架和区块链标准基础上开发的区块链客户端，将前一阶段获得的数据信息写入区块链，包括初始信息，如报价和聚合与匹配信息。在这个阶段，链上代码和链下处理的结合，也可以对区块链进行监管，保护相关人员的权利。 最后在结算阶段，区块链在收到买卖双方的匹配信息后，对交易进行结算。实现资产转接。这些交易过程是全程自动存储的，保证所有交易记录可以随时查取，并不被篡改。 特别是对能源交易来说，出现了能源币这种交易模式。那么能源币到底能不能取代传统货币呢？他安全吗？消费者信任他吗？说是你的隐私可以得到保护，但是在链上，别人如果都不知道你是谁，怎么知道买你的电。而且说是电力交易可追溯，具体的怎么对电力进行标注，这些都是问题。 在中国，传统消费者在中国还处于电卡交易制，就是往卡里充钱。是用多少交多少。对现在这种加密货币而言，普通人了解还没有那么多，所以很大的限制了他的发展。当然，未来也不排除加密货币的大规模使用，但是政策，环境，信任度，都是很大的挑战。
M2	我从事的行业主要是能源交易的农光互补这一块。更多的是偏向于价格机制这一块。我觉得中国政府在这一块做的还是很好的。 特别是对农民使用这种新能源的补贴很到位。在中国，输配电价由政府单独制定。交电费时的电价叫做“销售电价”。它由上网电价、输配电价、线损折价、政府基金及附加 4 部分组成。 农光互补光伏电站在不改变原有土地性质和地形地貌的基础上，因地制宜，将光伏发电与农业种植、畜牧有机结合，实现“板上光伏发电，板下现代农业”共同发展。在生态方面，具有显著的节能减排效益；在农业方面，创新农业发展新模式；同时，推动当地能源结构优化，创造良好的社会和经济效益。 但是一地两用，有人说好，也有人说不好。一方面能让作物获得足够光照的同时，保护作物不受冰雹、暴雨和猛烈光照的伤害，但是另一方面，这种光伏电板用旧了怎么

	办？怎么回收？会对环境造成污染吗？这种板的回收和销毁还是很浪费资源，造成浪费的。 农民现在觉得很好，因为有补贴，可以省钱。用不了的还可以卖回电网，双向收费。但是未来国家政策改变，是否会引起农民不再使用了呢，我们也不得而知。还有就是，有人说当他们看到自己家的储能不够的时候，就不在想交易了，他们怕电价会应声而涨。 此外，对消费者而言，我们经常会遇到的问题，比如说，“使用这个可以节省多少钱？”当然了，消费者最看重的还是经济利益。有的农户一个月花不了 10 块钱的电费。 他们肯定是考虑怎么省钱怎么来。 有顾客问“这个产品好用吗？“我一个老年人能操作的了吗？ 的确一个系统的好坏直接影响消费者想不想继续使用。但是现阶段，咱们还是和国家电网合办的项目，操作起来没有太大的障碍。一般人都可以掌控。 也有很少的人问“这真的是天然的吗？”， 你说呢？咱们用的是天然的自然的光。肯定是天然的。唯一不天然的就是这个光伏板，以后咱们应该怎么降解处理。现在通用的光板讲解需要 20 到 30 年，所以这对未来的环境发展有很大的压力。 所以，未来认为在能源交易这个问题上，最大的问题是政策问题。咱们现在更多的倡导节约能源，提倡使用清洁能源， 减少煤炭的燃烧和使用，减少 co2 的排放，从而保护环境。这个点对点交易，对于国内来说，是一件很新鲜的事情。只有从政策上发生改变，对能源保护和节约立法，才有可能促使交易的发生。
M3	首先能源架构牵扯到一个能源连不连接主网的问题。中国是要连接的。公司发的电，农民塑料大棚发的电，都要并到主网。然后再由主网输送给客户。现在的上的平台，基本上都是基于区块链技术的新平台。传统上，市场主体之间的交易是以集中的方式进行的：发电商和用户将数据上传到集中的处理系统，双方通过集中汇总或优化进行匹配。然而，随着多种类型和数量的分布式能源的出现，交易的数量、规模和信息数据都在增加，集中式的决策方式导致交易中心的运营成本提高，时间消耗延长。区块链技术作为一种去中心化的分布式记账系统，可以借助密码学原理解决交易过程中的所有权问题，有效解决交易过程中的信息不对称问题。 对市场来说。现阶段，发电点与工业用户之间的双边交易越来越广泛。个人之间的交易还是少。当然，在中国短时间内不可能，因为政府管控。在美国，这些都是可以的。几年前，美国推出一个社区的能源交易服务，好像叫做布鲁克林计划，就是在这个社区里，只要有安装这个太阳能光板，就可以加入这个计划。然后你就可以把你太阳光板上使用不完的电卖给你的邻居。但是貌似没有什么水花。 在中国，当然是不可能的。首先你要考虑安装问题。安在哪里啊。咱们到处是高楼。安装在农村。不过这个在中国也是有的。一种是农业光照，直接入网。国家每度电补贴 2 毛钱。另外一种是企业自己建立的光伏电厂，然后卖给一些大型机构，比如说学校，医院。 还有就是，企业间的相互买卖和个人用户之间的相互买卖是不同的。确切说，是有很大不同。对企业间的买卖来说，企业更期待得是是否够便宜。是不是能帮助他们节约成本。 其次他们考虑的是，能否真的保护环境。因为现在咱们国家的政策要求是企业要减排，所以他们的减排压力还是很大的。当然， 对于卖方而来说，则是更偏向于

	卖给大型企业，因为一次可以卖的多啊，电存放不住，普通民众也没有那么大的购买能力。 对用户来说，我自己本身认为，这种交易在短时间内实现是不现实的.或者说，在中国实现是不现实的。首先因为本身能源是稀缺物品，至少和铅笔来对比，他是稀缺产品，所以人们就需要去争取。也因此在中国，电力交易都是国家掌控的。是不允许个人之间随便买卖的。 同时由于电力市场的放开，消费者和生产者之间的这个界限就变得模糊了。这也就决定了电力交易类型和管理模式会发生改变。因此如果政府在未来大规模开放电力市场，即使做到和欧美一样，那可能也还是会有所不同。首先完全开放的市场，在中国还是行不通的。政府还是要有一定的把控。首先中国国情，意识形态都与西方不同。政府对市场秩序的混乱是很担心的。中国人真的很需要适度的管理。如果电在市场上随便卖卖，真的市场就会混乱。因此，即使中国未来开放电力市场，也是有选择开放。电力作为产品之间的交易，也会是很多大公司进行交易。个人可以根据自己的需求，在众多公司中选择。当然，这些公司肯定是信誉有保障的大公司。政府不会允许你让你个人进行交易。你自己房顶上产生的电能也只能卖给这些大公司或者国家。这些都是有记录，价格都是固定的，绝对不可能随便制定价格，改动市场秩序。 同时你提到西方国家市场上的产品琳琅满目。这个问题我不是很确定。这个琳琅满目是什么意思？电力的最终形式都是一样的。无论是太阳能光能，风能最终都会以电的形式被你使用。应该说，在西方市场上，你可以选择的公司更多。很多公司都有这个产品。其实说白了，还是你最终选择哪家公司的产品。只不过现在这些公司，都对自己的产品进行了重新包装。来自风电农场的，他们称天然环保无污染。来自他们自己公司光伏电站的可能是价格低廉最实惠。现在有个人或者社区内的交易，有可能你今天买的电视来自你儿子学校的屋顶上的，可能就是慈善助学了。 同时，消费者的欲望是无限的，人们总想更好。那么你所谓的定制能源可以解决什么问题呢。能满足人民日益增长的要求吗？能源的属性决定了，它在短时间内很难改变。那么它怎么去满足不断提高的需求呢。而且这种电力交易极大地取决于气候条件，比如说天气好的时候，储能高，天气不好的时候储能就低。那么天气的好坏就影响了交易的价格。 但是我们也不能排除一个问题。如果在一个小区里，普及率高了。那么很有可能，使用的人会越来越多。因为人们找到归属感了，周围的人都用了，自己是不是也要用？ 还有在很多有政策支持的农村，使用率也高。因为一个人用了，他的亲戚朋友也会用。你想你一个农村的亲戚来你家看你，看到你们家有安装这个光伏发电板，而且发电政府还有补助，所以你的亲戚就想也安装一个。同样的道理，朋友，邻居都是这么回事。而且由于农村人口居住的相对密集，所以一般一个村子里有人安装了，很快就会普及开来。当然前提是它要划算，如果化巨大的价钱来安装，这个问题就另当别论。 还要能源这个产品，是不是我们每天都要用，用的太频繁了，人们为什么还要去想今天用哪家能源？明天用哪家的能源呢？他们的最终属性有区别吗？最终都会以电的

	形式到你手里。那还会选择他到底是哪家的好吗？能源不是衣服，消费者不会考虑她穿上是否漂亮，考虑的就是偏不便宜？更高尚点说，能不能保护环境。 还有不得不提的是能源币是建立在区块链之上的支付方式，这种新的支付能否被大众接受。其实很多人是不接受的，现在国家也不断的提区块链这个问题。马上能源也要上链了。那下一步，就有能源币的使用。具体这个能源币怎么使用，和传统货币一样吗？当用户使用了能源币以后怎么交易？是使用电子钱包进行交易？还是预先充值，即使划取？这样操作安全吗？这期间有没有电子虚拟币的通膨？事实上，建立在以太坊基础上电子虚拟币是会随着以太坊的价格的升高和降低而出现价格的调整。但是能源的价格没有以太坊的波动幅度达，这就造成了一个差价。这个差价很多是不能被消费者接受的。但是如果我们使用法定货币，比如说美元，那么我们再使用虚拟货币的意义在哪里？从趋利避害的角度来看，虚拟货币在能源交易上应该走不远，或者说障碍很大。换句话说，虽然区块链最大的优点是建立智能合约，保护双方的隐私，但是相对于差价的损失，能否弥补保护隐私带来的好处还值得商榷。 曾经有客户就问我，区块链会断吗？其实我们不能笑他们，这也并不是无知的表现，并不是所有人都知道区块链，都知道数字货币。区块链是一种链式结构，其中的交易数据块按时间顺序链接。它本质上是一个去中心化的分布式数据库，通过非对称加密默克尔树等技术来保证数据的安全性，确保信息无法被外部攻击所篡改。但是虽然他们不知道什么是区块链，当我给他们普及以后，他们愿意去尝试了。因为他们觉得更安全了。这种顾客其实是占少数。大部分顾客，我给他们普及了，他们也不想用。因此对于这种新兴事物的接受和认识就需要一个阶段。就需要培养对它的信任度。但是这些问题被提出了，在未来的发展中，这些就都要考虑。我们现阶段，大家对区块链的态度还是趋之若鹜，根本不知道区块链在能源交易中最重要的是什么。
F1	现阶段新能源的交易在中国的发展很迅速，特别是基于区块链发展下的新能源交易，当然于我个人而言，我不会轻易使用，使用不适合中国的国情，你可以看一下，现在能源交易还处于公司与大型风力电厂之间的合作。是 B2B 的模式。个人用户之间是不倡导的，为什么，因为中国城市里到处高楼鳞次栉比，根本没地方安装太阳能光板，即使安装了，也无法为整栋楼里几百户同时供电。 至于电力平台构建，市场的话，目前我国新能源以保障性收购为主。与此同时，为促进新能源消纳，很多省区开展了新能源市场化交易探索，既包括开展大用户直接交易、发电权交易等中长期市场化交易，也建立了调峰辅助服务等短期市场。相比于中长期市场，短期市场能够更好的兼容新能源出力波动性与随机性的特点，对于促进新能源消纳具有重要作用. 对于现存的市场来说，如果大规模普及个人间的交易，可能并不乐观。因为在农村里，这个使用，国家是有补贴。如果农民要安装，国家是政策支持，资金鼓励。但是如果在城市里，首先不说有没有地方来安装这个太阳能光板， 只是这个初装费就会让很多人望而却步。很多人就会考虑，虽然新能源可以节约成本，但是这个安装费需要多少年回本？

	对未来来说，新能源与自备电厂置换、新能源参与大用户直供等新能源优先交易往往是年度电量交易，考虑新能源出力不确定、电力平衡困难等问题，需要进一步建立日前、日内等新能源优先交易机制。年度优先交易合同往往通过月度、日前等发电计划安排落实，可以在一定程度上为新能源让出电量空间，但是风电、光伏发电等发电出力具有随机性、间歇性等特点，发电量和实时的发电出力很难预测，年度、月度的发电计划在日前和实时落实时必须依靠电力系统所有环节的频繁、深度参与和协作。由于调峰补偿力度不够、需求侧响应机制缺失等因素，各类资源频繁参与系统调节的积极性不高，系统灵活性不足，影响新能源优先调度。一味依靠政策的调节方式难以有效调动发电机组、负荷等发挥最大调节潜力，需要进一步增加日前、实时的新能源短期交易，借助市场化手段挖掘系统灵活性 至于你说的能源互联技术，在中国现在很多专家，都在讨论，在广州和上海也有了试点。能源互联网是通过电子、电气和信息技术将大量分散的分布式能源连接起来，不仅实现了分布式能源的能源采集和接入互联网，以及局部区域内的电力传输和调配，还建立了各分布式能源之间以及分布式能源与集中式电网之间的互联，并通过大数据技术和信息化智能调控实现各种能源在整个网络中的互联。它还建立了分布式能源之间以及分布式能源与集中式电网之间的互连。 首先我先来回答你关于加密货币的问题。我只能现在的加密货币只是使用了区块链技术，未来能不能成为主流货币，这是存在很大的质疑的。加密货币和商品使用区块链交易还不一样。区块链交易让交易更安全了。但是加密货币不同。加密货币是有很大的风险的。加密货币是高风险的。它的波动性太大了。因此在交易之前，你就要做好准备，这种高风险的投资你是否可以接受。举个例子说，1 个比特币在 2011 年，可以买 1 个披萨。但是在 2020 年，它的价值是 5 万美元。这期间有多少投资回报率。千万不要认为这稳赚不赔，首先你即使买了比特币，它可能在你手里放不到这么长时间。你看到升值早就卖掉了。看到贬值，你更是恨不得马上出手。如果把比特币换成能源币，一个道理。现在的能源币大部分都是建立在以太坊基础上的。你需要用美元兑换以太坊币，进一步兑换成你所买的能源币。因此巨大的波动性可能导致你的能源币今天兑换 100 以太坊币，市场价格是 200 美元，但是只过了一天，它就只能值 50 美元了。因为以太坊币价值缩水了。所以一旦你知道了这个兑换机制，大部分会因为这个巨大的波动性，而不会去购买。至少不会大量购买。除非这个能源币能成为法定货币。因此加密货币由于波形过大，会带来金融风险。因此顾客不希望选择一些使用加密货币进行支付的产品。 未来，随着大规模可再生能源、电动汽车、储能设备和柔性负载的快速发展，分布式能源符合中国电力市场的发展需求和方向，并逐渐成为主要的发电来源，在中国具有广阔的应用前景。区块链技术作为一种新兴的去中心化数据管理技术，越来越受到各行各业的关注，但区块链在中国的电力应用还不成熟。因此，未来我们应该注意高比例分布式电源高效、经济使用的分布式电源优化，并通过虚拟电厂或需求响应的方式实现区域分布式电源的运行、管理和交易。在此基础上，应用区块链技术提高电力交易和管理的效率，增强区域分布式电源运行管理的自主性，提高运行的经济效益。
M4	中国的分布式发电这些年还是取得了很大的成效的。从分布式发电，到农光互补，再到渔光互补，现在也有户电光伏。发展还是很迅速的。 户用电站系统是在家庭住宅或附近建筑物上建设，利用光伏组件及系统将太阳光能直接转变为电能的新型发电方式。户用电站由光伏组件、逆变器、电表箱、监控模块、电缆、支架组成，组件将太阳能转换成电能，并网逆变器将直流电转换成交流电，电

	表箱对发电系统中的电能进行计量，监控系统方便业主随时关注电站系统发电情况。但是具体到用户之间的电力交易还不行，政策也不允许。 此外，9 号文的发布重启了中国电力市场改革;光伏、微网、储能技术成本逐年下降，分布式发电应用增长迅速。互联网的广泛应用以及能源数字化时代的到来，使的非传统能源企业开始侵入能源行业，改变着能源服务的内容和竞争局面。随着经济的放缓，用电需求增长放缓的同时电动汽车、能源管理系统等市场规模大幅增长，改变着电力需求与经济发展之间的关系。雾霾、污染，空气指数成了人 们生活中关注的对象，也成为社会媒体和政府关注的重点民生问题之一。 未来个性化的服务市场很大，人们需要针对自身，定制了个性化服务，消费者对个性化服务的期望也越来越高，但与此同时， 相关数据的安全性也成为消费者重点关注的问题。与此同时，使用者也考虑了这个交易系统的安全和稳定性问题。比如说一个消费者现在安装了这个交易系统。现在在有补贴的情况下，农民是有意愿持续使用的。但是当补贴变少，或者消失了，消费者是否还会继续使用，就要考虑一下。而且很多使用者反应，未来并不确定退出这个系统，如果退出，或者转换到原系统上，是否方便。 其次，绿色价值观的构建。咱们现在生活的地球是我们的，也不全是我们，是未来的子孙后代的。我们要绿色环保，保护我们的地球。 2016 年，在中国的十三五规划中，将区块链加入了国家信息规划。以后中国主要是要走协调发展，集成互补的能源互联的路子。现在，中国的能源结构正在由煤炭为主向多元化转变，能源发展势头由传统能源增长向新能源增长转变，清洁化、低碳化进程不断加快。近年来，风电、光伏、生物质能等可再生能源发展迅速，天然气的使用量不断增加，核电也在加快能源结构的转型。 传统的大型电源和电网的单一运行模式在一定程度上已不能满足当前能源转型的需要。因此，利用效率高、对环境负面影响小、提高能源供应可靠性、减少损耗、运行灵活、经济效益好的分布式能源正在迅速发展。天然气的分布式制冷、供热和供电、分布式光伏、分散式风电、分布式储能、需求方响应和分布式可再生能源供热等技术的发展，导致了能源转型的加速。 再次，信息分享与资源共享。也就是一直说的用户间能源交易。从全球范围来看，消费者的信息分享和资源共享使用户可以有更大的议价 能力，也使他们无需购买资产却仍可以享受服务，虽然在能源消费过程中，这一趋势尚未凸显，但消费者也意识到，通过信息分享和资源共享，在充分竞争的市场上，可将个人议价能力提升为“集合议价能力”，从而提升收益或降低成本。 同时，用户“热爱分享”的趋势，使社交网络平台成为能源供应商的有力工具。企业通过社群平台推出特别优惠、忠诚度计划及类似活动，还可利用这些渠道收集消费者反馈，了解其需求和预期，从而评判消费者情绪，识别有关媒体和客户服务问题。
F2	首先我在英国一家电力比价公司工作。 和 go compare 一样，我们的主要工作是根据客户的主要需求，为用户选择合适的能源公司。今天我就来讲一下，点对点能源消费背后的深层动机吧。 首先我想强调的是不为个人客户服务的。我们服务的都是公司。至于为什么选择公司，主要是个人消费对于这种新鲜的新能源的交易还不普及。他们

	更倾向的是使用便宜的套餐。对于个人来说，白天家里都没有人，他们更希望在使用峰谷的的时间段，获得更低的价格。比如说晚上洗衣服，如果电费够低的话，很多人会选择的。比如说，现在的 British gas 等大公司，都是很愿意推出这种套餐的。 还有就是，其实英国人还是很喜欢尝试新的产品的。主要你的产品价格够低。所以经常有客户转换自己的家庭供电公司。他们会经常比来比去，看哪家更便宜。 相比公司来说，公司的消费动机是很复杂的。 公司不止单纯的为了节省开支。我们的客户有很多考虑的不是成本上的节省，他们考虑的更多的能否为公司带来公司文化上的契合。举个例子， 有些公司文化倡导的就是绿色环保，所以他们在电力公司的选择上更愿意使用清洁能源。 对于这种新的清洁能源的交易，他们也更愿意尝试。因为这是他们公司文化里标榜的。 对于个人，我就不知道文化有没有区别的了，个人一般都是为了节省开支吧。当然环保主义者还是从环保主义出发的，相信这种产品一旦普及，很多环保人士还是愿意尝试的。 据我所知，英国现在正在对个人点对点交易模式进行试点。其实挑战是有的，我们都知道英国的光照分布很不均匀。一年有 200 天在下雨，风倒是很大。南部和北部的日照时长也有很大的区别。所以对于南部来说，这个项目可能更有前景。对于日照时间短的地方，用户自己发的电都不知道能不能维持供给，更不用说交易了。 但是怎么说呢，英国的社区发展还是很完善的，所有社区里如果普及率大的话，邻居之间的交易应该是很有市场的。而且英国对居民的隐私保护做得很好。 而且现在很多英国的公司都在尝试这种点对点的新能源交易。但是主要还是以商用为主。UCL 和 eon 有一个合作项目，是专门针对新型住宅小区的。这个的优势也是建立在新小区的基础上。一方面这样会增加售卖，另一方面测试小区内用户交易的情况。因此，在未来，如果有大量的公司进入这个市场。那么这很快就和现在市场差不多。零售还是渠道，都会有，主要定位的消费人群不同。当可持续能源和传统能源一样，可以在市场上进行交易，那么新能源就是一种再简单不过的产品。也就是你对不同能源公司进行选择的过程。归根结底还是选择与公司呈现出来的相关信息。这个公司的声誉怎么样？这个公司的服务质量好吗？我们打客服电话，问题能得到有效解决吗。因此，你在网上浏览这个公司的信息，看大家对他的评价怎么样。如果大家一直认为这个公司很好，那它提供的产品也是值得信赖的。那么很多顾客会对它进行选择。而且如果你一直使用 British gas， 你觉得这个公司的产品很好，当这个公司相继推出新能源交易的产品，你也会继续选择使用。 我对中国的能源模式不是很清楚。据我所知， 中国的能源应该是国有的吧？所以进行能源交易可行吗？有法律允许吗？在英国，所有的电通过国家电网早出来以后，会有不同的公司来销售，所以点对点交易的模式可行。 对于英国来说，市场前景还是有的。英国对这一块立法管理做得不错。英国有种能源使用评估证书 EPC， 这个是根据你使用的能源情况，对你的房子，无论是个人住宅和公司住宅进行的用电评估。这个 EPC 直接与地税相关，EPC 越高，地税越多。其实这一定侧面上反映的也是你的用电习惯,以及你对 CO_2 减排做出的贡献。你有没有随手关灯，随手关掉电脑，使用 led 节能灯泡，都会通过 EPC 反映出来。现在英国对 EPC 这一块管理的特别严格，如果这个房子超过 band D, 这个房子就很难出租。政府要强制你进行电力改造。所以说英国的对能源保护，和环境保护的法规还是很健全的，因此说这个能源交易在英国来说是有前景的。但是与此同时，我们也可以从另一个方面来考虑着这个问题. 有些人是从内心节省能源，无论出发点是经济考量，还是环保考量。有些人就不在乎能源的使用，自己需要多少就用多少。因此，在未来的能源交

	易中，面对这两种群体，提供的交易选择对象就会不同。比如说 EPC 标准低的，可能更倾向于环保产品。而不在乎的群体可能更关注电力的供给是否稳定。 对于未来来说，有两个问题需要考虑。一是初装费用的问题。这个初装费用可是真的不低。在英国，一般的太阳能公司会先来你们家进行实际测量。即使你在网上填取了相关信息。那也只是初期的评估。当他们获取了你的地址信息以后，就会和你联系，实际上门测量。包括你们家几间卧室，要安装几块板。然后计算出安装的费用。这个费用包括太阳能光板的费用和安装费。现在的技术，太阳能光板有了大幅度的提升，但是储能和续航还是很大问题。因此，很多人在初期得到太阳能光板的报价的时候就放弃了。即使太阳能公司给你一些相应的折扣，但是还是有很多人放弃了。还有一个有意思的事情就是，来咨询这个问题的很多都是中年人，他们的消费能力能够支撑他们进行尝试。但是话句话说，如果这个项目由政府进行帮助，结果可能会大大不同。 二是人们更关心的到底是买的谁家的电。你刚才也说，区块链可以标记，但是区块链不是实物，怎么对电的流向进行标记，咱们不知道。 但是相信，如果真的可以标记，那可能会激发人们使用的热情。 因为球迷可能想买它支持的球队的主场旁边的太阳能或者风能发电厂的电。 三是，到底这种产品是不是绿色的。怎么标记这种绿色的来源。能源最终产生的形式都一样。当然，区块链是一种好的标记方式。但是具体怎么标记，可能需要长期的给顾客普及知识。因为大部分人对区块链还是陌生的。

English version

Thank you for coming here today. More and more energy consumers are now turning into prosumers. Thanks to innovations in distributed energy resources including photovoltaic cells and wind turbines, users are using energy and producing it at the same time. This is the most fundamental motivation for creating peer-to-peer energy trading platforms. I am conducting a study to understand how customers perform on products for P2P energy trading. I have a few questions for you and thank you for your answers.

Can you introduce the P2P energy trading project based on own field please?

• Yes

• No

iv. Can you describe how does current p2p energy set up/ price mechanism/ market mechanism/ information system/regulation works?

v. What do you think are the main challenges in the current energy setting up/ price mechanism/ market mechanism/ information system/regulation?

vi. Are there any challenges and opportunities for the future P2P energy trading?

This is all the questions I would like to ask you. Thank you so much for your participation. If you have any questions would like to ask me, please feel free to send me an e-email: shan.shan@northumbria.ac.uk.

Thanks again!

Translation of Interviews

PROFILE	DATA
M1	About the current construction of China's energy trading grid, it is mainly based on agricultural use of electricity, with the main grid connected. Users can either be self-sufficient or sell to the main grid in reverse. As far as energy trading itself is concerned, the focus of different departments must be different. For the power operation department, the concern is the local consumption of electricity and how to reduce the power loss in the transportation process. Also, we need to consider that new energy sources are inherently unstable; there is a big difference between having the sun and not having the sun, and a big difference between having the wind and not having the wind. The difference between the sun and no sun, wind, and no wind, is also huge, so when they are integrated into the grid, they cause a lot of volatility to the grid itself. Therefore, many customers ask us if they are connected to the grid when they first use them. That is, are they related to the main grid? Because they are afraid that they will not be able to generate enough electricity for themselves and that they will have to take the risk. Not being connected to the mains network makes them feel that the electricity supply is unstable. Also, the connection to the main grid is promoted and required by the state to earn additional subsidies for electricity. Many farmers are very frugal with electricity and are basically self-sufficient. As China is a vast country, there is a big gap between the complementary east and west photoelectricity, and in these parts of Ningxia, the electricity generated simply cannot be used up. That's why there is west-east electricity transmission and west-east gas transmission. For the market, there are now three main types of energy trading in China. One is agriculture and solar power. That is, the government has certain financial subsidies for farmers who use solar power. Another is a blockchain distributed ledger technology and artificial intelligence supported by solar energy and battery storage-based electricity trading platform. German companies, for example, and a Dutch company, distro, are also doing research in this area. You can look it up. In this case, consumers have a better idea of where the energy is going. Buyers and sellers are thus informed about market preferences, supply and demand and energy price dynamics, which in turn leads to more accessible and more efficient electricity sales and purchases. These platforms are now mainly in transition to blockchain-based platforms. Each market participant is equipped with an artificial intelligence instrument that acts as a "trading agent", essentially removing the middleman, saving time and costs, and facilitating transactions. In addition, blockchain can effectively reduce the cost of trust, improve the efficiency of transactions, safe and reliable, and the decentralisation, open sharing, autonomy and security

	of distributed energy, in line with the development trend of energy. The application of blockchain in the energy sector will effectively support the open interconnection of multiple types of energy systems and the extensive and in-depth participation of multiple users. By jointly maintaining a trusted distributed ledger, the effective connection of energy flow, information flow and value chain in future energy transactions can be realised. For users, their enthusiasm for using it is still high because of the state subsidies, friendly policies, and the savings they make. Then there are the large power generation farms, such as wind power plants, which are usually established in the suburbs with a large plot of land, and after collecting the electricity, sell it to some companies to achieve two-way benefits. In foreign countries, many companies are now joining blockchain platforms. Through blockchain, they can trade with their neighbours and communities. Let's not analyse whether blockchain is good or not, but nowadays all products are on the chain. It is said that it can achieve anonymity and more security. An Australian company, for example, has launched a blockchain-based system for trading surplus household energy in urban areas. When an electricity producer enters and sells green electricity to the distribution network, the distribution network operator confirms this, and upon successful confirmation, a smart contract generates the corresponding virtual currency and offers it to the green electricity producer as a reward. The whole transaction is simple, and trust is increased using virtual currencies. Some companies have also introduced energy coins, which are tokens based on Ethereum, or bitcoin. Customers are not very receptive to this model of trading. Because energy coins fluctuate so much? and with traditional dollar settlements, they don't seem to benefit. Today bitcoin is $40,000, tomorrow it could be $1,000 longer. So, this type of trading is not reciprocal for energy itself. Further, cryptocurrencies are not regulated by governments or central banks. This is troubling. Under the existing political ideology in China, there is certainly no way that the currency could lose the regulation of the central government. The Chinese are smart enough to believe that if regulation is deregulated, there will be an endless stream of illegal transactions by then. Of course, I don't think Western societies will legalise such cryptocurrencies either. Although the ideology of Western countries is different from that of China, the goal is the same. Western countries may treat cryptocurrencies differently, and they may see it as a commodity. It is possible to look at this commodity with an open mind. But in China, this cryptocurrency poses a threat to our monetary system, it threatens the functional position of the national bank. So, it is not possible for cryptocurrencies to be legal tender in China. And it is not possible to trade freely in energy coins. Unless the billion yuan is settled directly. Or direct settlement with electronic money introduced by the central bank. Some foreign platforms are now offering some special packages, such as installing solar panels and getting a corresponding amount of energy coins for free or getting a discount on your next purchase. In fact, these are not very tempting for consumers. The first thing

	consumers consider is who do I buy from and who do I sell to, and is it reliable? Can I get a refund if I don't use it? Energy, unlike other products, is a hard requirement. Consumers don't care what kind of electricity you have, at the end of the day it's all the same electricity anyway, you can fry and cook and watch TV. And can this discount be compared to the installation cost; can you get an effective return in a few years. There is an age limit on solar panel light panels, which generally need to be replaced in 20 years. The subsequent problem then is that you must invest in new costs to renew the light panels again. But the key is whether the proposed investment has paid for itself in those 20 years. If you live in a place with enough sunlight to have plenty of sunshine all year round, that could be considered. So, this project is going to be a bit longer lasting in Australia. If it was the UK where it rains 200 days a year, it probably wouldn't be suitable. Then again, when customers buy something, they will consider the reputation of the brand, its image, and all sorts of other things before they slowly develop loyalty. Energy is one thing. Many companies are doing this selling of sustainable energy. Then there are customers who pay particular attention to the reputation of the company, to see if it is a big company and if it is trustworthy. There is also the fact that consumers are more concerned about the level of penetration, if my neighbours have installed it, do I want to install it, so that it is better integrated in a community. In other words, if the household penetration rate in a community is high, consumers are more likely to find a sense of belonging. Then again, you say your best friend tells you that they have installed the product in their home and can trade with others. Would you want to install it too? Would you be even more motivated if he then told you what a great deal it was? What are the challenges now, mainly in terms of setting up this platform for peer-to-peer trading, this end-to-end energy trading involves a lot of communication and information at all stages of negotiation, execution and clearing? For both centralised and distributed end-to-end energy trading markets, a large amount of bilateral communication is required between the market coordinator and individual users; for decentralised markets, a large amount of communication is also required between individual users. For future research, communication load and investment requirements could be an important dimension in evaluating end-to-end energy trading market mechanisms and trading platforms. The impact of non-ideal communication conditions, such as communication delays, distortions and failures, on end-to-end energy trading needs to be further assessed. In addition, the optimal configuration and operation of communication networks for end-to-end energy trading needs to be further investigated. Also, in fact many end-to-end prices are even higher than market prices. Some studies suggest that people's preference for self-sufficiency will reduce the likelihood of end-to-end energy trading, as they believe that this makes people inclined to give a higher price than their real value when pricing their own local generation. In other words, one's own electricity consumption is met, and one realises one's control over the electricity and the user's expectations of the actual value of the product are not as high. So, what we call the economic

	benefits that the user values most may not be valid. Also, if you are in a community where none of your neighbours use it, I believe it is unlikely that you will use it yourself. Because you also must take the risk, the risk of unstable electricity. For energy policy, this is something that I believe China is doing very well. Our country is very supportive. From the 13th Five-Year Plan to now, the country has done a good job with agriculture and fisheries, and with complementary light energy generation. Specifically in foreign countries, we are not the same as foreign countries. In foreign countries, electricity is privatised, but we are nationalised. We are nationalised. We are nationalised. So, in the future, if the electricity is traded between neighbours, there should be a long way to go. And if you look at all the tall buildings in China now, there's no room for solar panels. Putting up a few small solar panels wouldn't be enough for the whole building's inhabitants. As for the future development, I would like to mention blockchain. Blockchain is very hot in China now, and everyone is talking about blockchain, and about being on the chain. How exactly does blockchain work? Generally, in the information dissemination and aggregation stage, qualified electricity users and distributed energy generators register and make requests to purchase and sell electricity through a website or application, and the users reach a transaction intention at the application layer, i.e., the matching of buyers and sellers is completed at this stage without matching in the blockchain. Information is then transmitted to the blockchain. The blockchain client, developed based on the application framework and blockchain standards, writes the data information obtained in the previous phase to the blockchain, including initial information, such as offers and aggregation and matching information. The combination of on-chain code and off-chain processing at this stage also allows the blockchain to be supervised and the rights of those involved to be protected. Finally, in the settlement phase, the blockchain settles the transaction after receiving the matching information from the buyer and seller. The transfer of assets is achieved. These transaction processes are stored automatically throughout, ensuring that all transaction records can be accessed at any time and are not tampered with. For energy trading in particular, a trading model such as energy coins has emerged. So, can energy coins replace traditional currencies or not? Is he safe? Do consumers trust him? It says your privacy can be protected, but on the chain, how does anyone else know to buy your electricity if they don't even know who you are. And it says that electricity transactions can be traced, specifically how to label the electricity, these are the questions. In China, traditional consumers in China are still in the electricity card transaction system, which is to charge money to the card. It's pay as much as you use. For this kind of cryptocurrency now, ordinary people don't know that much yet, so it greatly limits his development. Of course, the mass use of cryptocurrencies is not ruled out in the future, but the policy, environment and trust are all big challenges.
M2	I work in an industry that focuses on the agricultural and solar complementary aspect of energy trading. It is more on the price mechanism side. I think the Chinese government has done a very good job in this area. In particular, the subsidies for farmers to use this new energy source are very good. In China, transmission and distribution prices are set by the government

alone. The price of electricity is called the "sales tariff" when you pay for it. It is made up of four parts: the feed-in tariff, the transmission and distribution tariff, the line loss discount, and government funds and surcharges.

Based on not changing the nature of the original land and topography, the PV power station can combine PV power generation with agricultural planting and animal husbandry to achieve "PV power generation on the board and modern agriculture under the board". In terms of ecology, it has significant energy saving and emission reduction benefits; in terms of agriculture, it innovates a new model of agricultural development; at the same time, it promotes the optimisation of the local energy structure and creates good social and economic benefits.

But two uses for one land, some say yes, others say no. On the one hand, it allows crops to get enough light while protecting them from hail, heavy rain, and fierce light, but on the other hand, what about this kind of photovoltaic panels when they are old? How can they be recycled? Will they pollute the environment? The recycling and destruction of such panels is still a waste of resources and causes waste.

Farmers now think it is good because they have subsidies and can save money. Those that can't be used can also be sold back to the grid and charged in both directions. But whether a change in national policy in the future will cause farmers to stop using them, we don't know. There is also the fact that some people say that when they see that they don't have enough storage in their homes, they are not thinking about trading, they are afraid that the price of electricity will rise in response.

Also, for consumers, we often get questions like, "How much money can I save by using this?" Of course, it is the financial benefits that are most important to consumers. There are farmers who spend less than $10 a month on electricity. They are thinking about how they can save money.

Some customers ask "Does this product work? "Can I operate it as an elderly person? It is true that the goodness of a system directly affects whether consumers want to continue using it or not. But at this stage, we are still a joint project with the National Grid, so there are no major barriers to operation. It is manageable for the average person.

There are few people who ask, "Is it really natural?" What do you think? We are using natural light. It must be natural. The only thing that is not natural is the photovoltaic panels and how we should dispose of them in the future. Nowadays it takes 20 to 30 years to explain the common light panels, so this puts a lot of pressure on the future development of the environment.

So, in the future, I think the biggest problem in this issue of energy trading is the policy issue. We now advocate more energy conservation, advocate the use of clean energy, reduce the burning, and use of coal, reduce co2 emissions, so as to protect the environment. This peer-

	to-peer trading is a very new thing for the country. Only a change in policy and legislation on energy conservation and saving will make it possible for trading to happen.
M3	First of all, the energy architecture involves a question of whether energy is connected to the main network or not. China is to be connected. The electricity generated by companies, the electricity generated by farmers' plastic sheds, must be connected to the main grid. Then the main grid is used to deliver electricity to customers. The platforms that are on now are basically new platforms based on blockchain technology. Traditionally, transactions between market players have been conducted in a centralised manner: generators and customers upload data to a centralised processing system and the two parties are matched through centralised aggregation or optimisation. However, with the emergence of multiple types and volumes of distributed energy sources, the number, size and information data of transactions are increasing, and the centralised approach to decision-making is leading to higher operating costs and longer time consumption for trading centres. Blockchain technology, as a decentralised distributed bookkeeping system, can solve the ownership problem in the transaction process with the help of cryptographic principles and effectively address the information asymmetry in the transaction process. For the market. At this stage, bilateral transactions between power generation sites and industrial users are becoming more and more widespread. Transactions between individuals are still rare. Of course, this is not possible in China any time soon because of government control. In the US, these are possible. A few years ago, a community-based energy trading service was launched in the US, something like the Brooklyn Project, which is a scheme that you can join in this community if you have this solar panel installed. You can then sell the electricity you don't use on your solar panels to your neighbours. But there doesn't seem to be any waterworks. In China, of course, it's not possible. First you must think about the installation. Where to install it. We have tall buildings everywhere. Install it in the countryside. But this is also available in China. One is agricultural light, which goes directly to the grid. The state subsidises 20 cents per kWh. The other kind is photovoltaic power plants set up by companies themselves, which are then sold to large institutions, such as schools and hospitals. There is also the fact that the mutual sale between enterprises is different from the mutual sale between individual customers. To be precise, there is a big difference. In the case of business-to-business sales, the business is more interested in whether it is cheap enough. Whether it will help them to save money. The second thing they look for is whether they can really protect the environment. They are under a lot of pressure to reduce emissions because of our national policy. Of course, for the seller, the preference is to sell to large enterprises, because they can sell more at a time, the electricity cannot be stored, and the public does not have that much purchasing power. For the user, I believe that such a transaction is not realistic in the short term. Or rather, it is not realistic in China. First, because energy is a scarce commodity, at least compared to

	pencils, it is a scarce product, so people need to fight for it. That's why in China, electricity trading is controlled by the state. It is not allowed to be bought and sold between individuals. At the same time, because of the liberalisation of the electricity market, the line between consumer and producer has become blurred. This has led to a change in the type of electricity trading and the mode of management. So, if the government were to open the electricity market on a large scale in the future, even if it were to do so in the same way as in Europe and the US, it would probably still be different. A completely open market, for a start, would still not work in China. The government still must have some control over it. First, China's national conditions and ideology are different from the West. The government is very worried about the disorder in the market. The Chinese really need a moderate amount of regulation. If electricity is sold and sold casually in the market, really the market will be chaotic. Therefore, even if China opens its electricity market in the future, it will do so selectively. Electricity as a product will be traded between many large companies as well. Individuals will be able to choose between the many companies according to their needs. Of course, these companies must be large companies with a guaranteed reputation. The government will not allow you to let you trade personally. The electricity you generate on your own roof can only be sold to these large companies or to the state. These are recorded and the prices are fixed. It is impossible to just set prices and alter the market order. At the same time, you mention that there is a wide range of products on the market in Western countries. I'm not sure about that. What do you mean by a wide range of products? The final form of electricity is the same. Whether it's solar light, wind energy, you end up using it in the form of electricity. I should say that in the Western market you have a much wider choice of companies. A lot of companies have this product. It really comes down to which company's product you end up choosing. It's just that these companies, now, have repackaged their products. The ones from wind farms, they claim are natural and environmentally friendly and non-polluting. The ones from their own company photovoltaic plants are probably the most affordable at low prices. There are now personal or intra-community deals, and it's possible that the TV you buy today from the roof of your son's school could be a charity aid. At the same time, consumer desire is infinite, people always want better. So, what problem can you solve with what you call bespoke energy. Will it meet the growing demands of the people? The nature of energy dictates that it is difficult to change in a short period of time. So how does it go about meeting the ever-increasing demand. And this kind of electricity trading greatly depends on weather conditions, for example when the weather is good the storage is high and when the weather is bad the storage is low. Then the weather affects the price of the transaction. But we can't rule out a problem either. If the penetration is high in a small community. Then most likely, more and more people will use it. Because people have found a sense of belonging, and everyone around them is using it, should they not use it too? In many rural areas, where there is a policy of support, there is also a high rate of use. Because if one person uses it, his relatives and friends will also use it. If a relative from a rural area

	comes to visit you and sees that you have a photovoltaic panel installed in your home and that the government is subsidising the generation of electricity, then your relative will want to install one too. The same goes for friends, neighbours and so on. And since rural areas are relatively densely populated, it's usually common for people in a village to install one. Of course, the prerequisite is that it is cost effective, but if it costs a fortune to install it, this is a different matter. We also need think about energy, do we use it every day, so often that people have to think about which energy source to use today? Which energy source will be used tomorrow? Is there a difference in their final properties? Ultimately, they will all come to you in the form of electricity. So, would one still choose which one is better for him? Energy is not a garment, the consumer does not consider whether she looks good in it or not, all he considers is whether it is cheap or not? To put it more nobly, will it protect the environment? It is also important to mention that Energy Coin is a payment method built on the blockchain, whether this new payment can be accepted by the public. In fact, many people don't accept it, and now the country keeps mentioning blockchain as an issue. Soon energy is going to be on the chain too. Then, the next step is the use of energy coins. How will this energy coin be used? Is it the same as traditional currency? How do users trade with the energy coins after using them? Will they use an e-wallet for transactions? Or is it pre-funded and withdrawn? Is it safe to do so? Is there any inflation of electronic virtual coins in the meantime? The fact is that e-virtual coins based on Ethereum are subject to price adjustments as the price of Ethereum goes up and down. However, the price of energy does not fluctuate as much as Ether, which creates a price differential. Much of this price difference is not acceptable to consumers. But if we use a fiat currency, say the US dollar, then what is the point of using a virtual currency again? From a tendency to avoid harm, virtual currencies should not go far in energy trading, or there are significant obstacles. In other words, while the biggest advantage of blockchain is the creation of smart contracts to protect the privacy of both parties, it is debatable whether the benefits of protecting privacy can be compensated for relative to the loss of price difference. A client once asked me if the blockchain would break. We can't laugh at them, and it's not a sign of ignorance that not everyone knows about blockchain and digital currencies. A blockchain is a chained structure in which blocks of transaction data are linked in chronological order. It is essentially a decentralised distributed database that secures data through technologies such as asymmetric encryption Merkle trees, ensuring that information cannot be tampered with by external attacks. But although they didn't know what blockchain was, when I popularised it for them, they were willing to try it. Because they feel more secure. This kind of customer is in the minority. Most of the customers, even if I popularise it, they don't want to use it. So, it takes a while for this new thing to be accepted and understood. There is a need to develop a level of trust in it. But these issues are being raised, and in the future, these will have to be considered. We are still, at this stage, in a rush to blockchain and have no idea what the most important thing about blockchain in energy trading is.

F1	At this stage, new energy trading is developing rapidly in China, especially new energy trading based on blockchain development. Of course, for me personally, I would not use it easily, using it is not suitable for China's national conditions, you can look at it, now energy trading is still in the cooperation between the company and large wind power plants. It's a B2B model. It is not advocated between individual users. Why? Because there is no room to install solar panels in Chinese cities where tall buildings are lined up everywhere, and even if they were, they would not be able to supply power to hundreds of households in the whole building at the same time. As for the construction of the electricity platform, the market, then, is currently China's new energy to guarantee the acquisition of the main. At the same time, to promote new energy consumption, many provinces, and regions to carry out new energy market-based trading exploration, including both to carry out direct transactions with large users, power generation rights trading and other medium and long-term market-based transactions, but also to establish a short-term market such as peak-adjustment auxiliary services. Compared to the medium and long-term market, the short-term market is more compatible with the volatility and randomness of new energy output and plays an important role in promoting the consumption of new energy. For the existing market, it may not be optimistic to popularise inter-individual trading on a large scale. This is because in rural areas, the use of this, the state is subsidised. If a farmer wants to install it, the state supports it with policies and financial incentives. But in the city, first, not to mention the availability of space to install the solar panels, just the initial installation fee will make many people discouraged. Many people consider that although new energy can save costs, how many years will it take to pay back the installation costs? For the future, new energy and captive power plant replacement, new energy participation in direct supply to large users and other new energy priority trading is often annual power trading, considering the uncertainty of new energy output, power balancing difficulties and other issues, need to further establish a new energy priority trading mechanism such as day-ahead, intra-day. Annual priority trading contracts are often implemented through monthly, day-ahead and other power generation plan arrangements, which can to a certain extent give room for new energy to generate power, but wind power and photovoltaic power generation output is random and intermittent, and it is difficult to predict power generation and real-time power generation output, so annual and monthly power generation plans must rely on frequent and deep participation and cooperation of all parts of the power system when implemented in day-ahead and real-time. collaboration. Due to factors such as insufficient compensation for peaking and the lack of demand-side response mechanisms, all types of resources are not highly motivated to participate in system regulation frequently, and the system is not flexible enough, which affects the priority dispatch of new energy. It is difficult to effectively mobilise generating units and loads to maximise their regulation potential by relying on policy-based regulation, and there is a need to further increase short-term trading of new energy resources

on a day-to-day and real-time basis, and to tap into system flexibility through market-based means.

As for the energy interconnection technology you mentioned, in China many experts, are now discussing it, and there are pilot projects in Guangzhou and Shanghai. The energy internet is a way of connecting a large number of decentralised distributed energy sources through electronic, electrical and information technology, not only to achieve energy harvesting and access to the internet for distributed energy sources, as well as power transmission and deployment within a local area, but also to establish interconnection between each distributed energy source and between distributed energy sources and the centralised grid, and to achieve the interconnection of various energy sources throughout the network through big data technology and information-based intelligent regulation and control. It also establishes the interconnection between distributed energy sources and between distributed energy sources and the centralised grid and realises the interconnection of various energy sources across the network through big data technology and information-based intelligent regulation. It also establishes the interconnection between distributed energy sources and between distributed energy sources and the centralised grid.

Let me start by answering your question about cryptocurrencies. I can only use blockchain technology for cryptocurrencies now, and there is a big question mark over whether they will become a mainstream currency in the future. Cryptocurrencies and commodities using blockchain transactions are not yet the same. Blockchain transactions have made transactions more secure. But cryptocurrencies are different. Cryptocurrencies are very risky. Cryptocurrency is high-risk. It's too volatile. So, before you trade, you need to be prepared for whether this high-risk investment is acceptable to you. Let's say, for example, that 1 bitcoin could buy 1 pizza in 2011. But in 2020, it's worth $50,000. What is the return on investment in that period? Don't ever think it's a sure thing. First, even if you buy bitcoin, it probably won't sit in your hands for that long. You'd have sold it by now if you saw the appreciation. When you see it depreciate, you'll be even more desperate to get rid of it right away. If you replace bitcoin with energy coins, one thing applies. Most of the energy coins now are built on Ethereum. You need to exchange USD for Ether coins, which are further converted into the energy coins you buy. So, the huge volatility could result in your energy coin being exchanged for 100 Ether coins today at a market price of $200, but after only one day it will only be worth $50. This is because the value of the Ether coin has shrunk. So, once you know this exchange mechanism, most will not buy because of this huge volatility. At least not in large quantities. Unless this energy coin can become legal tender.

In the future, with the rapid development of large-scale renewable energy, electric vehicles, energy storage devices and flexible loads, distributed energy meets the development needs and direction of China's electricity market, and gradually becomes the main source of power generation, with broad application prospects in China. Blockchain technology, as an emerging decentralised data management technology, is receiving increasing attention from various industries, but blockchain is not yet mature for power applications in China. Therefore, in the future, we should pay attention to distributed power optimization for efficient and economical

	use of high proportion distributed power, and realise the operation, management, and transaction of regional distributed power by means of virtual power plants or demand response. On this basis, blockchain technology should be applied to improve the efficiency of power trading and management, enhance the autonomy of regional distributed power supply operation and management, and improve the economic efficiency of operation.
M4	China's distributed power generation has still achieved great results over the years. From distributed power generation, to agricultural photovoltaic, to fishery photovoltaic, and now also household photovoltaic. The development is still very rapid. Household power station system is a new type of power generation built on the family residence or nearby buildings, using photovoltaic modules and systems to directly transform the sun's energy into electricity. The household power station consists of photovoltaic modules, inverters, meter boxes, monitoring modules, cables, and brackets. The modules convert solar energy into electricity, the grid-connected inverter converts direct current into alternating current, the meter box measures the electricity in the power generation system and the monitoring system facilitates the owner to keep an eye on the power generation situation of the power station system. But specifically, electricity trading between users is not yet possible, and the policy does not allow it. In addition, the release of No. 9 restarted the reform of China's electricity market; the cost of photovoltaic, microgrid and energy storage technologies is decreasing year by year, and the application of distributed power generation is growing rapidly. The widespread use of the internet and the advent of the digital age of energy have led to the intrusion of non-traditional energy companies into the energy industry, changing the content and competition of energy services. As the economy slows, the growth in demand for electricity slows while the market for electric vehicles and energy management systems grows significantly, changing the relationship between electricity demand and economic development. Haze, pollution and air indices have become a concern in people's lives and one of the key livelihood issues that social media and governments are focusing on. There is a large market for personalised services in the future and people need to tailor them to themselves. Consumers' expectations of personalised services are growing, but at the same time, the security of the relevant data has become a key concern for consumers. At the same time, users are also considering the security and stability of the transaction system. Let's say that a consumer has now installed this transaction system. Now, with subsidies, farmers are willing to continue using it. But when the subsidy becomes less, or disappears, it is a matter of consideration whether the consumer will continue to use it. And many users respond that they are not sure about withdrawing from the system in the future, and whether it would be easy to do so, or to switch to the original system. Secondly, the construction of green values. The planet we live on now is ours, and not all of us, it is the future generations. We need to go green and protect our planet. In 2016, in China's 13th Five-Year Plan, blockchain was added to the national information plan. In future, China will mainly follow the path of coordinated development and integrated and complementary energy interconnection. Now, China's energy structure is changing from coal-based to diversified, the momentum of energy development is changing from traditional energy growth to new energy growth, and the process of cleanliness and low carbonization is accelerating. In recent years, renewable energy sources such as wind power, photovoltaic and

	biomass have developed rapidly, the use of natural gas is increasing, and nuclear power is also accelerating the transformation of the energy mix. The traditional single mode of operation of large power sources and power grids can no longer meet the needs of the current energy transition to a certain extent. As a result, distributed energy sources with high utilisation efficiency, low negative impact on the environment, improved reliability of energy supply, reduced losses, flexible operation, and good economic efficiency are developing rapidly. The development of technologies such as distributed cooling, heating and power supply for natural gas, distributed photovoltaics, decentralised wind power, distributed energy storage, demand-side response and distributed renewable energy for heating has led to an acceleration of the energy transition. Again, information sharing and resource sharing. This is what has been referred to as inter-user energy trading. Globally, information sharing and resource sharing by consumers allows for greater bargaining power and access to services without the need to purchase assets, and although this trend has not yet been highlighted in the energy consumption process, consumers are also aware that through information sharing and resource sharing, in a fully competitive market, individual bargaining power can be increased to " collective bargaining power" in a fully competitive market, thereby increasing revenues or reducing costs. At the same time, the "sharing" trend of users has made social networking platforms a powerful tool for energy suppliers. Companies using social platforms to launch special offers, loyalty programmes and similar activities can also use these channels to gather consumer feedback, understand their needs and expectations, gauge consumer sentiment and identify issues related to media and customer service.
F2	Firstly, I work for an electricity price comparison company in the UK. Like go compare, our main job is to select the right energy company for our customers based on their main needs. Today I'll talk about the deeper motivations behind peer-to-peer energy consumption. The first thing I want to emphasise is that we don't serve individual customers. We serve companies. As to why we choose companies, the main reason is that individual consumers are not yet comfortable with the idea of trading in this new energy source. They are more inclined to use cheaper packages. For individuals, who have no one at home during the day, they prefer to get a lower price during the peak and valley hours of use. Many people will choose to do their laundry at night, for example, if the electricity bill is low enough. Large companies such as British gas, for example, are now very willing to offer such packages. There is also the fact that British people like to try new products. The main thing is that the price of your product is low enough. So, there are often customers who switch their home electricity company. They will often compare to see which one is cheaper. The motivation for companies to spend money is very complex compared to companies. Companies are more than simply trying to save money. Many of our customers are not thinking about cost savings, they are thinking about whether they can bring a company culture to the table. For example, some companies have a green culture, so they prefer to use clean energy in their choice of power company. They are also more willing to experiment with new clean energy deals. Because that's what their culture says. For individuals, I don't know if there's a difference in culture, individuals generally do it to save money, I guess. Of course, environmentalists are still environmentalists, and I believe that many of them will be willing to try these products once they become popular.

	I understand that the UK is now piloting a personal peer-to-peer trading model. There are challenges, and we all know that the UK has a very uneven distribution of light. It rains 200 days of the year, and the wind pours. There is also a big difference between the hours of sunlight in the south and the north. So, for the south, the project is probably more promising. For places with short sunshine hours, it's not clear that customers can maintain their supply with the electricity they generate themselves, let alone trade it. But what can be said, the UK is still very well developed in terms of communities and there should be a market for trading between neighbours in all communities if the penetration rate is high. And the UK does a good job of protecting the privacy of its residents. Many companies in the UK are now experimenting with peer-to-peer energy trading, but mainly for commercial use. UCL and eon have a joint project for new housing estates, but mainly for commercial use. The advantage of this is that it is also based on new communities. On the one hand this will increase the sales and on the other hand test the user transactions in the small communities. In the future, therefore, if many companies enter this market. Then this will soon be like the current market. Retail or channel, both will be there, mainly targeting different consumer groups. When sustainable energy can be traded on the market as well as conventional energy, then new energy could not be a simpler product. It is also the process by which you make a choice between different energy companies. Ultimately the choice is related to the information presented by the company. What is the reputation of this company? Is the quality of service from this company good? Do we call the customer service number, and do we get effective solutions to our problems? So, you browse the information about this company online and see what people say about him. If people consistently think that this company is good, then it offers a trustworthy product. Then many customers will make a choice about it. And if you've been using British gas, and you think the company's products are good, you'll continue to choose to use them when the company comes out with new energy deals one after another. I don't know much about the energy model in China. As far as I know, energy in China is state-owned, right? So, is it feasible to trade energy? Is there a law that allows it? In the UK, all the electricity comes out of the national grid early on and is sold by different companies, so the peer-to-peer trading model works. For the UK, the market is still promising. The UK has done a good job of regulating this piece of legislation. The UK has an Energy Performance Assessment Certificate (EPC), which is an assessment of the electricity used in your home, whether it is a personal home or a company home, based on the energy you use. This EPC is directly linked to land tax and the higher the EPC, the higher the land tax. In fact, it is also a reflection of your electricity consumption habits and your contribution to CO_2 reduction. Whether you turn off your lights, turn off your computer or use led energy saving light bulbs, this is reflected in your EPC. In the UK, EPCs are very strict, and it is very difficult to rent a house if it exceeds band D. The government must force you to carry out electrical renovations. So, the UK has very strong regulations on energy conservation and environmental protection, so there is a future for energy trading in the UK. But at the same time, there is another way to look at this issue. Some people are conserving energy from the inside, whether it is from an economic or environmental point of view. Some people don't care about energy use, they use as much as they need. Therefore, in the future, the energy trade will offer different options to these two groups. Those with low

	EPC standards, for example, may be more inclined to environmentally friendly products. And the group that doesn't care may be more concerned about the stability of the supply of electricity. For the future, there are two issues to consider. One is the issue of initial installation costs. This initial installation cost is not low. In the UK, the average solar company will come to your home first to take physical measurements. Even if you have filled in the information online. That is only the initial assessment. Once they have your address information, they will contact you to take a physical measurement. This includes how many bedrooms in your home and how many panels to be installed. Then calculate the cost of the installation. This cost includes the cost of the solar panels and the installation fee. With today's technology, solar light panels have improved dramatically, but energy storage and range are still big issues. As a result, many people give up when they get an initial quote for solar light panels. Even if the solar companies give you some discounts accordingly, there are still many people who give up. Another interesting thing is that many of the people who come to ask about this are middle-aged people whose spending power can support them in trying it out. But then again, if this project had been helped by the government, the result could have been much different. Secondly, people are more concerned about whose electricity they are actually buying. You said earlier that blockchain can be tagged, but blockchain is not a physical object, so how to tag the flow of electricity, we don't know. But I believe that if it really could be tagged, that might inspire people to use it. Because a fan might want to buy electricity from the solar or wind power plant next to the home stadium of the team it supports. Thirdly, whether this product is green or not. How to mark the source of that green. The energy is ultimately generated in the same form. Blockchain is, of course, a good way to tag it. But it may take a long time to spread knowledge to customers about how to mark it. Because most people are still unfamiliar with blockchain.

Appendix B: Survey

分布式能源交易平台客户问卷 (CHINESE VERSION)

非常感谢您同意参与本次调查。本次研究工作的目的是了解客户对使用 P2P 能源交易产品的态度意向。问卷共分五部分，完成时间约需 20 分钟。
您在问卷中提供的信息将仅用于研究目的，而不会用于商业目的。本研究产生的结果将为人类与技术互动研究的文献做出贡献。通过本次调查，您也将获得一些关于 p2p 能源交易的知识。
您是在没有受到胁迫或任何压力的情况下自愿参与本次调查的。
您可以随时撤回您的许可，我们鼓励您尽可能诚实地回答问题。
在整个数据收集过程中，您的匿名权和保密权将受到保护。
您可以随时查阅资料，并能与研究者联系。
本人完全理解以上声明，并同意参加本次调查。 [单选题] *

○是

○否

背景知识

点对点能源交易(P2P 能源交易)是指两个或多个并网方之间的能源买卖。通常以太阳能的形式，任何多余的能源都可以通过一个安全的平台转让和出售给其他用户。点对点能源交易让消费者可以选择决定向谁购电，向谁卖电。
随着新能源分布式发电项目(Peer to Peer 能源交易)的快速发展，如太阳能交易、风能交易与 B2B、B2Cand C2C。另外，近期各国对分布式发电采取了不同的政策，如创新中心与传统能源企业的联合研究。众所周知，分布式发电是在用电地点或邻近用电地点，进行高电压传输。与集中式发电相比，它具有降低电能损耗、节约输电成本、减少土地资源占用等优点。除此之外，分布式发电的目标是实现能源的灵活性。
目前，大部分的交易项目都是基于区块链技术的。在分布式账本技术的帮助下，P2P 交易平台可以高效地工作. 在 P2P 交易项目中，区块链技术可以帮助降低收益者之间电力交易的交易成本。所有的用户和消费者的信息都可以通过区块链的智能合约得到保护。 而所有的交易都可以用能源币或货币进行交易（能源币可以兑换成比特币）。
不过，在您正式开始调查之前，我们还是要让您清楚地了解项目的优势和劣势，以及不同产品的区别，以帮助您做出决策。
项目的优势包括。
1. 能源独立，意味着你可以购买和出售你想要的能源。
2. 环境保护
3. 节省费用
4. 创新，因为区块链是一项新技术

项目的缺点包括
1. 能源供应不稳定。受天气影响较大。
2. 能源并入主网后，电压稳定性将受到影响。
3. 利用可再生能源交易，需要 10 年内安装费用才能回本。
4. 如何处理废旧电池板是个问题，因为它们很难被降解。但随着技术的发展，情况变得越来越好。

产品信息介绍 （五种）
零售平台： 会以较低的价格为竞争提供了绝对优势。例如，电力用户可以在批发市场上进行直接和廉价的购买。
供应商平台：旨在帮助每一个卖家提高其产品的价值。例如说通过电池储能和光伏板的创新，提供特定的电池系统和光伏板供应商，可以减少与高压电力有关的功能损失。
社区平台：建立在自己居住社区里的交易平台。可以方便的与邻里进行交易。
区块链平台：是基于区块链技术设计的。区块链是一种涉及透明度和易用性的创新技术，可以毫不费力地产生和消耗能源。通过使用基于信用的支付系统，也可以提高交易速度和效率，这将消除区块链在 P2P 能源交易中使用可靠中介的需要。
混合平台：由社区平台、储能电池和国家电网共同整合形成的平台。这种类型的平台为稳定的能源供应提供了更大的系统兼容性，为电网运营商提供了更大的可预测性。

个人信息
1. 性别 [单选题] *

○男

○女

○选择不说

2. 年龄 [单选题] *

○18-24

○25-34

○35-44

○45-54

○55-64

○64+

3. 教育程度 [单选题] *

○高中

○本科

○硕士

○博士

4. 家庭总收入(年) 如果未婚，请计算父母在内 [单选题] *

○¥50000-99999

○¥100000-199999

○¥200000-299999

○¥300000-399999

○¥400000-499999

○¥500000+

5. 婚姻状态 [单选题] *

○已婚

○未婚

○离异

6. 家庭成员 [单选题] *

○1-3 人

○4-6 人

○7+人

7. 您有多少子女 [单选题] *

○0

○1

○2-4

○4+

○选择不说

（二） 房屋信息

8. 您居住的区域 [单选题] *

○城市

○农村

9. 您现在居住的住房 [单选题] *

○小于 50 平方米

○50-99 平方米

○100-149 平方米

○150-199 平方米

○200-249 平方米

○250-300 平方米

○大于 300 平方米

10. 您现在居住的房子的居住状态 [单选题] *

○自有房（无贷款）

○自有房（有贷款）

○经济适用房

○租房

○宅基地自建房

○合租房

○其他

11. 您现在房子的房龄 [单选题] *

○少于 5 年

○5-9 年

○10-19 年

○20-30 年

○多于 30 年

12. 您现在居住房子的窗户玻璃是 [单选题] *

○单层玻璃

○双层玻璃

（三）能源使用习惯

13. 您日常交通工具（选择最常用的一种） [单选题] *

○公共交通（公共汽车，地铁，轻轨）

○自驾

○自行车

○出租车

○火车

○共享单车

○拼车

14. 您会在外出时关灯吗？ [单选题] *

○会

○有时候会

○不会

15. 你会在外出时关闭电脑吗？ [单选题] *

○会

○有时候会

○不会

16. 你出门的时候会不会把暖气开着？ [单选题] *

○会

○有时候会

○不会

17. 你会在 40 度以下洗衣服吗？ [单选题] *

○会

○有时候会

○不会

18. 不需要给手机充电时，你会把手机充电器插在插座上吗？ [单选题] *

○会

○有时候会

○不会

19. 你会用烧水壶烧超过需要的水量吗？ [单选题] *

○会

○有时候会

○不会

20. 您更愿意跟谁进行能源交易 [单选题] *

○认识的人

○不认识的人

○不在乎

21. 您希望以什么样的支付方式来进行能源交易？ [单选题] *

○纸质货币

○电子货币

22. 我没有认真完成问卷 [单选题] *

○是

○否

（四） 能源交易产品选择行为

请仔细阅读以下陈述。在您的评分中使用以下量表，说明您的同意/不同意程度（1-7；1=非常不同意，7=非常同意）

23. 我是独一无二的人
24. 他人的话不能影响我的决定
25. 我可以承担我自己做出的决定
26. 我觉得很快参与 P2P 能源交易是一种个人义务
27. 我觉得我有道德义务尽快参与 P2P 能源交易
28. 我觉得 P2P 能源交易对未来一代和我们生活的世界是有益的。
29. 我认为能源系统的变化是受欢迎的
30. 我认为能源交易中的不确定性是可以接受的
31. 我认为能源的公平分配很重要
32. 我希望能源交易的过程是有序和稳定的
33. 我认为政府有必要加强对交易过程的监管
34. 我认为政府有必要加强对平台的监管
35. 我认为能源交易网站/应用程序所提供的信息是可靠的。
36. 我认为能源交易网站/应用程序提供的信息与实际情况相差无几。
37. 我相信 P2P 能源交易平台产生的信息。
38. 我知道我的福利受到关注
39. 当我向我的服务提供商倾诉我的问题时，我知道他们会以理解的态度作出回应。
40. 我可以指望我的服务提供者考虑他们的行为对我的影响。
41. 切换系统的便利性对我很重要
42. 关闭我的账户对我来说很重要
43. 容易取回余额很重要
44. 更多的电池储存可以鼓励我进行 P2P 能源交易
45. 稳定的供应可以鼓励我进行 P2P 能源交易。
46. 集成在小型电网内可以帮助我帮助我进行 P2P 能源交易
47. 为我省钱
48. 有助于降低我的电费
49. 使我在经济上受益。
50. 我可以接受因无法将能源币兑换成传统货币，或无法以合理价格兑换的损失
51. 我可以接受因对手方未能履行合同付款或结算义务而造成的损失，这一点很重要。
52. 我可以接受由于市场波动造成的损失
53. 有助于减少对环境的不利影响
54. 有助于减少能源和其他资源的消耗
55. 使我能够成为一个更有社会责任感的文明人。
56. 使用能源令牌让我感到安全
57. 使用加密货币不会让人担心
58. 基于比特币的能源代币没有风险
59. 从别人那里购买能源听起来很刺激
60. 购买邻居的能源的想法是令人愉快的
61. p2p 能源交易很有趣
62. 该供应商以诚信著称
63. 该供应商以关注客户而闻名
64. 该供应商在市场上有不良声誉

65. 你有多大意愿去寻求新奇的信息？
66. 你是否会寻找新奇和不同的东西？
67. 你有多愿意尝试新的交易类型？
68. P2P 能源交易使我能够获得不同的来源
69. 如果我选择与他人进行交易，可以获得广泛的供应商
70. 帮助我与他人联系
71. 让我能够认识附近的人
72. 让我发展社会关系。
73. 当你办公室/教室里的大多数人都使用 P2P 能源交易时，你会不会使用 P2P 能源交易？
74. 当你所在社区的大多数人都使用 P2P 能源交易时，你会使用 P2P 能源交易吗？
75. 我周围的朋友越是开始使用 P2P 能源交易服务，我就越是愿意使用 P2P 能源交易服务。
76. 使用 P2P 能源交易让我成为志同道合者群体的一员
77. 使用 P2P 能源交易让我成为兴趣相投的群体的一部分
78. 大多数对我很重要的人都会认为，使用 P2 P 能源交易是一个明智的想法
79. 大多数对我重要的人都会认为我应该使用 P2 P 能源交易
80. 对我很重要的家人会认为使用 P2 P 能源交易是一个明智的想法

（五）. 产品信息

81. 请从下面能源交易平台中选择你需要平台

[ ]集成平台

[ ]零售平台

[ ]自动售卖平台

[ ]区块链平台

[ ]社区平台

English version

Thank you very much for agreeing to participant in this survey. The purpose of this research work is to understand customers’ attitude intention to use P2P energy trading. The whole questionnaire contains five sections and will takes you approximately 20 minutes to complete.

The information provided by you in this questionnaire will be used for research purpose only and not for commercial purpose. The results generated by this study will contribute to the literature in the human-technology interactions studies. You will also get some knowledge about p2p energy trading through this survey.

You participated in this survey voluntarily without coercion or under any pressure.

You can withdraw your permission at any time and are encouraged to be honest as possible with your answers.

Your right to anonymity and confidentiality will be protected during the whole process of data collection.

You can access the information and are able to contact with the researcher at any time.
If you are understanding the above statements, please click the right box shown in the follows:

o I am totally understand above statements and agree to join in this survey

o I do not want to join in this survey Thank you so much!

Background

Peer to Peer energy trading (P2P energy trading) is the buying and selling of energy between two or more grid-connected parties. Often in the form of solar energy, any excess energy can be transferred and sold to other users via a secure platform. Peer-to-peer energy trading allows consumers to decide to purchase electricity from whom, and who they sell it to.

With the rapid development of new energy distributed power generation projects (Peer to Peer energy trading), such as solar energy trading, wind energy trading with B2B, B2Cand C2C. Also, recently, different countries took different policies towards to the distributed power generation, for example, the joint research between innovation centre and traditional energy corporate. As we know, distributed generation is located at or adjacent to the place of power consumption, which elevated voltage transmission. Compared to centralized power generation, it has the advantage of reducing power losses, saving transmission costs, and reducing the use of land resources. Besides of this, the goal of distributed generation is to achieve the energy flexibility.

Currently, most of the trading projects are based on the blockchain technology. The P2P trading platform can work efficiently with the help of distributed ledger technologies, of which the most prominent type is blockchain. Blockchain technology can help reduce the transaction costs for electricity trading among prosumers in a P2P trading scheme. All the prosumers and consumers information can be protected by smart contract with blockchain. And all the transaction can be deal with either energy coin or currency (energy coin can be exchanged to bitcoin).

However, before you formally start the survey, we still need to let you clearly understand the advantages and disadvantages of the projects.to help you make the decisions.
The advantages of the project include,

1. Energy independent means you can buy and sell energy as you want
2. Environment protection
3. Saving cost
4. Innovation as blockchain is a new technology

The disadvantages of the project include

1. Instability of the energy supply. It is significantly impacted by the weather.
2. The volage stability will be impacted when the energy integrated into the main grids,
3. The installation cost can be returned at least in 10 years by using renewable energy trading.
4. How to deal with the wasted panel and wind turbs is a problem, as they are difficult to be degraded. But the situation became better and better with the technology development.

Section 1: Demographic Information
Nationalities ____

1. Gender:

○Male

○Female

○Not to answer

2. Age

○18-24

○25-34

○35-44

○45-54

○55-64

○64+

3. Education

○High school/GED

○College

○Bachelor's degree

○Master's degree

○Doctoral degree

4. Household Income per year (if not married, family include parents)

○Under £9999

○£10000-£19999

○£20000-£29999

○£30000-£39999

○£40000-£49999

○£50000-£59999

○ Above £60000

5. Marital status

○Single

○Married

○Divorcee

○Prefer not to say

6. Family member

○1-3

○4-6

○7+

7. How many children do you have?

○ None

○ 1

○ 2-4

○ More than 4

Section 2: Building information

8. Living area

○Urban

○Rural

9. Bedrooms

○1 bedroom

○2 bedrooms

○3 bedrooms

○4 bedrooms

○4+ bedrooms

10. Is the house or flat in which you live (single code)

a) Owned outright

b) Owned with a mortgage or loan

c) Rented from the council

d) Rented from a housing association

e) Rented from someone else (e.g. a private landlord)

f) Other

11. Building age

○ under 5years

○ 5 to 9 years

○ 10 to 19 years

○ 20 to 30 years

○30 years above

12. What is your window glazing

○ single glazing

○ double glazing

Section 3: Energy consumption information

13. Daily Travel (Choose one mostly used)

○ Public transport

○ Self-driving

○ Recycling

○ Walk

○ car sharing

○ Taxi

14. Would you turn off the lights when going out?

○ every time

○ sometimes

○ never

15. Would you shut down your computer when going out

○ every time

○ sometimes

○ never

16. Would you leave the heating on when you go out

○ every time

○ sometimes

○ never

17. Would you wash the clothes at 40 degrees or less

○ every time

○ sometimes

○ never

18. Would you leave the mobile phone charger switched on at the socket when not in use

○ every time

○ sometimes

○ never

19. Would you boil the kettle with more water than you are going to use

○ every time

○ sometimes

○ never

20. Who would like to trade your energy with? *

○People you know

○People you don't know

○I don't care

21. What kind of payment method would you like to use for energy trading? [ONE CHOICE] *

○Local Currency

○Cryptocurrency

22. Please select ‘yellow’ in the statement.

○ yellow ○red ○black ○ pink ○brown

Section 4：Please use the following scale in your rating to indicate how you agree/disagree (1-7; 1=strongly disagree, 7=strongly agree)

INDIVILISM

23. I am unique person

24. What the other said cannot influence my decision

25. I can afford the decision made my self

COLLECTIVISM

26.I feel that participating inP2P energy trading soon is a personal obligation

27. I feel that I have a moral obligation to participate in P2P energy trading soon

28.I feel P2P energy trading is benefit for the future generation and the world we live.

LIBERALISM

29. I think the energy system changes are welcome

30. I think uncertainty in energy trading is acceptable

31. I think fair distribution of energy is important

CONSTRUCTIVISM

32. I want the process of energy trading to be orderly and stable

33. I think it is necessary for the government to strengthen regulation of the transaction process

34. I think it is necessary for the government to strengthen regulation of the platform

INFORMATION QUALITY

35. I think the information provided by the energy trading website/app is reliable.

36. I think the information provided by the energy trading website/app is closed to the real situation.

37. I trust the information generated by P2P energy trading platform.

SERVICE QUALITY

38. I am aware of my welfare is concerned

39. When I confide my problem to my service provider, I know they will respond with understanding

40. I can count on my services provider considering how their actions affect me

SYSTEM QUALITY

41. Easy to switch the system is important to me

42. Easy to close my account is important to me

43. Easy to get the balance back is important

STABLE SUPPLY

44. More battery stored can encourage me make P2P energy trading

45. Stable supply can encourage me make P2P energy trading**.**

46. Integrated within the min grids can help me help me make P2P energy trading

ECONOMICAL BENEFIT

47. Saves me money

48. Helps lower my electricity bill

49. Benefits me financially.

FINANCIAL LOSSES

50. I can accept losses due to inability to convert energy coins to conventional currencies, or not at a reasonable price

51. I can accept losses due to counterparties failing to meet contractual payments or settlement obligations is important

52. I can accept losses due to market volatility

SUSTAINABLITY

53. Helps reduce the negative impacts on the environment

54. Helps reduce the consumption of energy and other resources

55. Allows me to a more socially responsible civilization.

FEELING OF UNCERTAINTY

56. Using energy token make me feel security

57. Using cryptocurrency will not make worried

58. Energy token based on bitcoin has no risk

ENJOYMENT

59. Buying energy from others sounds exciting

60. The idea pf purchasing neighbours' energy is enjoyable

61. P2P energy trading is fun

REPUTATION

62. This supplier has a reputation for being honest

63. This supplier is known to be concern about customer

64. This supplier had a bad reputation on the market

NOVELTY SEEKING

65. How willing are you to seek out novel information?

66. Do you search for the new and different?

67. How willing are you to try new type of trading?

VARIETY SEEKING

68. P2P energy trading allows me to get access to different sources

69. Wide ranges of supplier are available if I choose to trade things with others

SOCIAL BENEFIT

70. Helps me connect with others

71. Allows me to get to know people from the neighbourhoods

72. Allows me to develop social relationships.

SOCIAL INFLUENCE

73. Will you use P2P energy trading when most people in your office/classroom use P2P energy trading

74. Will you use P2P energy trading when most people in your community use p2p energy

75. The more my friends around me start using P2P energy trading services, the more I am willing to use P2P energy trading service. COMMUNITY BELONGING 76. The use of P2P energy trading allows me to be part of a group of like-minded people 77. The use of P2P energy trading allows me to be part of a group of people with similar interests SUBJECTIVE NORMS 78. Most people who are important to me would think that using the P2 P energy trading is a wise idea 79. Most people who are important to me would think I should use P2 P energy trading 80. My family who are important to me would think that using P2 P energy trading is a wise idea

Section 5: Selection of Products
81. Please select the platform you need from the following energy trading platforms [ ] Hybrid platform (renewable energy trading schedules integrated with the main grid and community based). [ ] Retail platform [ ] Big brand vending platform [ ] Blockchain platform [ ] Community platform

Appendix C: Factor Measurement

Construct	Reference	Measures
1 INDIVILISM	(Ianole-Călin et al., 2020)	23. I am unique person 24. What the other said cannot influence my decision 25. I can afford the decision made my self
2 COLLECTIVISMS	(Ianole-Călin et al., 2020)	26.I feel that participating inP2P energy trading soon is a personal obligation 27. I feel that I have a moral obligation to participate in P2P energy trading soon

		28. I feel P2P energy trading is benefit for the future generation and the world we live.
3 LIBERALISMS	(Caldwell et al., 2020)	29. I think the energy system changes are welcome 30. I think uncertainty in energy trading is acceptable 31. I think fair distribution of energy is important
4 CONSTRUCTIVISM	(Caldwell et al., 2020)	32. I want the process of energy trading to be orderly and stable 33. I think it is necessary for the government to strengthen regulation of the transaction process 34. I think it is necessary for the government to strengthen regulation of the platform
5 INFORMATION QUALITY	(Seddon, 1997)	35. I think the information provided by the energy trading website/app is reliable. 36. I think the information provided by the energy trading website/app is closed to the real situation. 37. I trust the information generated by P2P energy trading platform.
6. SERVICE QUALITY	(Wang & Tang, 2003)	38. I am aware of my welfare is concerned 39. When I confide my problem to my service provider, I know they will respond with understanding 40. I can count on my services provider considering how their actions affect me
7. SYSTEM QUALITY	(Wang & Tang, 2003)	41. Easy to switch the system is important to me 42. easy to close my account is important to me 43. easy to get the balance back is important

8. STABLE SUPPLY	(Hahnel et al., 2020; Wang & Sun, 2017)	44. More battery stored can encourage me make P2P energy trading 45. Stable supply can encourage me make P2P energy trading. 46. Integrated within the min grids can help me help me make P2P energy trading
9. ECONOMICAL BENEFITS	(Bock et al., 2005)	47. Saves me money 48. Helps lower my electricity bill 49. Benefits me financially.
10. FINANCIAL LOSSES	(Abramova & Böhme, 2016)	50. I can accept losses due to inability to convert energy coins to conventional currencies, or not at a reasonable price 51. I can accept losses due to counterparties failing to meet contractual payments or settlement obligations is important 52. I can accept losses due to market volatility
11. SUSTAINABLITY	(Hamari et al., 2016)	53. Helps reduce the negative impacts on the environment 54. Helps reduce the consumption of energy and other resources 55. Allows me to a more socially responsible civilization.
12. FEELING OF UNCERTAINTY	(Kankanhalli et al., 2005)	56. Using energy token make me feel security 57. Using cryptocurrency will not make worried 58. Energy token based on bitcoin has no risk
13. ENJOYMENTS	(Hamari et al., 2016)	59. Buying energy from others sounds exciting 60. The idea pf purchasing neighbours' energy is enjoyable 61. P2P energy trading is fun
14. REPUTATIONS	(Doney & Cannon, 1997)	62. This supplier has a reputation for being honest 63. This supplier is known to be concern about customer

		64. This supplier had a bad reputation on the market
15. NOVELTY SEEKING	(Hirschman, 1980)	65. How willing are you to seek out novel information? 66. Do you search for the new and different? 67. How willing are you to try new type of trading?
16. VARIETY SEEKING	(Kim & Jin, 2020)	68. P2P energy trading allows me to get access to different sources 69. Wide ranges of supplier are available if I choose to trade things with others
17. SOCIAL BENEFIT	(Lin & Bhattacherjee, 2008)	70. Helps me connect with others 71. Allows me to get to know people from the neighbourhoods 72. Allows me to develop social relationships.
18. SOCIAL INFLUENCE	(Bearden et al., 1989)	73. Will you use P2P energy trading when most people in your office/classroom use P2P energy trading 74. Will you use P2P energy trading when most people in your community use p2p energy 75. The more my friends around me start using P2P energy trading services, the more I am willing to use P2P energy trading service.
19. COMMUNITY BELONGING	(Möhlmann, 2015)	76. The use of P2P energy trading allows me to be part of a group of like-minded people 77. The use of P2P energy trading allows me to be part of a group of people with similar interests
20 SUBJECTIVE NORMS	(Bock et al., 2005)	78. Most people who are important to me would think that using the P2 P energy trading is a wise idea 79. Most people who are important to me would think I should use P2 P energy trading

		80. My family who are important to me would think that using P2 P energy trading is a wise idea

Appendix D: Informed Consent Form for Research Participants

Faculty of Environment and Engineering

Northumbria University
NEWCASTLE

Informed Consent Form for research participants

Title of Study	A Machine Learning Approach to Solving the Complete Cold Start Problem in Recommendation Systems: Building the P2P Energy Trading Recommendation System with Theory of Consumption Values
Person(s) conducting the research:	Shan Shan
Programme of study:	Computer and Information PGR
Address of the researcher for correspondence:	Room 108, Ellison B Newcastle Upon Tyne United Kingdom, NE1 2JA
Telephone:	07588008746
E-mail:	Shan.shan@northumbria.ac.uk
Description of the broad nature of the research:	This research is part of my doctoral study. The research is for academic purpose only and not for commercial purpose. The purpose of this research work is to mitigate the cold start problem toward p2p energy trading recommendation system. This research includes 2 phase studies. It begins with two qualitative studies for exploring critical factors that may significantly influence trading behaviour and following up with a quantitative study to collect a large sample so that the researcher can generalize results to a population. And then, in the second phase, a machine learning approach is applied by the comparison of different algorithms.

	This research will contribute on creating a contribution to build the new p2p energy trading recommendation datasets to fill the cold start gap in the energy trading recommendations system and provide accurate personal recommendations toward p2p energy trading. The findings generated from this research also can help R&D managers in p2p energy trading market to find the most effective products in the market.
Description of the involvement expected of participants including the broad nature of questions to be answered or events to be observed or activities to be undertaken, and the expected time commitment:	1. Participation in this survey is voluntary without coercion or under any pressure. 2. Participants can withdraw their permission at any time and are encouraged to be honest as possible with their answers. 3. Participants can access the information and are able to contact with the researcher at any time. 4. The expected interview time is an hour for each participant. 5. The questionnaire will be posted on a Chinese and western market research survey website for 3 weeks.
Description of how the data you provide will be securely stored and/or destroyed upon completion of the project.	1. To protect participants' right to anonymity and confidentiality, a coding system will be adopted to identify the participants instead of using their real name or personal ID. 2. The data will be password- protected and only can be assessed by a researcher. 3. Hard copy of the questionnaire is not required. Electronic records will be stored in logical files structures and indexed using logical file. 4. The expected time of storing the data is approximate the length of completion of project adds to 5 years. Arrangements for the archiving of electronic materials will be made within the Environment and Engineering Faculty.

Information obtained in this study, including this consent form, will be kept strictly confidential (i.e., will not be passed to others) and anonymous (i.e., individuals and organisations will not be identified *unless this is expressly excluded in the details given above*).

Data obtained through this research may be reproduced and published in a variety of forms and for a variety of audiences related to the broad nature of the research detailed above. It will not be used for purposes other than those outlined above without your permission.

Participation is entirely voluntary, and participants may withdraw at any time.

By signing this consent form, you are indicating that you fully understand the above information and agree to participate in this study based on the above information.

Participant's signature: Date: Date:

Appendix E-Participant Debrief

Participant code:14207

PARTICIPANT DEBRIEF

Name of Researcher: Shan Shan
Name of Supervisor (if relevant): Dr Honglei Li

Project Title: A Machine Learning Approach to Solving the Complete Cold Start Problem in Recommendation Systems: Building the P2P Energy Trading Recommendation System with Theory of Consumption Values

1. What was the purpose of the project?

   The purpose of this research work is to mitigate the cold start problem toward p2p energy trading recommendation system. This research includes 2 phase studies. It begins with two qualitative studies for exploring critical factors that may significantly influence trading behaviour and following up with a quantitative study to collect a large sample so that the researcher can generalize results to a population. And then, in the second phase, a machine learning approach is applied by the comparison of different algorithms.

   This research will contribute on creating a contribution to build the new p2p energy trading recommendation datasets to fill the cold start gap in the energy trading recommendations system and provide accurate personal recommendations toward p2p energy trading. The findings generated from this research also can help R&D managers in p2p energy trading market to find the most effective products in the market.

2. How will I find out about the results?

The data will be analysed approximately 3 weeks after taking part of the interview. The final study will be completed on 01/03/2021. The researcher will email you a general summary of the results if you would like to know.

3. If I change my mind and wish to withdraw the information I have provided, how do I do this?

If you wish to withdraw your data, then please email the researcher named in the information sheet within 1 month of taking part and given me the code number that was allocated to you (this can be found on your debrief sheet). After this time, it might not be possible to withdraw your data as it could already have been analysed.

The data collected in this study may also be published in scientific journals or presented at conferences. Information and data gathered during this research study will only be available to the research team identified in the information sheet. Should the research be presented or published in any form, all data will be anonymous (i.e., your personal information or data will not be identifiable).

All information and data gathered during this research will be stored in line with the Data Protection Act and will be destroyed 60 months (the length of completion of the research adds to 5 years) following the conclusion of the study. If the research is published in a scientific journal, it may be kept for longer before being destroyed. During that time the data may be used by members of the research team only for purposes appropriate to the research questions, but at no point will your personal information or data be revealed. Insurance companies and employers will not be given any individual's personal information, nor any data provided by them, and nor will we allow access to the police, security services, social services, relatives, or lawyers, unless forced to do so by the courts.

If you wish to receive feedback about the findings of this research study, then please contact the researcher at shan.shan@northumbria.ac.uk

This study and its protocol have received full ethical approval from Faculty of Environment and engineering Ethics Committee. If you require confirmation of this, or if you have any concerns or worries concerning this research, or if you wish to register a complaint, please contact the Chair of this Committee, stating the title of the research project and the name of the researcher.

Thanks again for your participation.