Explainable Online Recommendation Systems with Self-Identity Theory and Attribute Learning Method

Zequn Li

PhD

2019

Explainable Online Recommendation Systems with Self-Identity Theory and Attribute Learning Method

Zequn Li

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 2019

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 11/11/2016.

I declare that the Word Count of this Thesis is 36187 words.

I declare that parts of the following papers and works have been included in this thesis.

1. Li, Zequn, Honglei Li, and Ling Shao. ”Improving Online Customer Shopping Experience with Computer Vision and Machine Learning Methods.” In International Conference on HCI in Business, Government, and Organizations, pp. 427-436. Springer, Cham, 2016. (Chapter 2,3)
2. Li, Zequn, Honglei Li, and Ling Shao. ”Improving Attribute Classification with Imperfect Pairwise Constraints” (2019). ICEB 2019 Proceedings. (Chapter 4)
3. Li, Honglei; Li, Zequn; and Tian, Jing, ”The Design of an Online Roommate Finding System for Property Management Portals—the Virtual Community Perspective” (2018). ICEB 2018 Proceedings. 4. And associate Industry project RoomMate Matching System funded by Nicelandlords Newcastle (Chapter 7)
4. Wen, Zhenyu, Rawaa Qasha, Zequn Li, Rajiv Ranjan, Paul Watson, and Alexander Romanovsky. ”Dynamically partitioning workflow over federated clouds for optimising the monetary cost and handling run-time failures.” IEEE Transactions on Cloud Computing (2016). (Chapter 7)

Name: Zequn Li

Signature:

Date: 23/01/2019

Acknowledgements

I would like to thank my supervisor Dr. Honglei Li for her guidance in research, support in life and encouragement to go through my difficulties. I would also like to thank my second supervisor Prof. Ling Shao and Dr. Longzhi Yang, without their support I cannot accomplish my research.

I would also like to thank the Northumbria University for providing opportunity and scholarship to support me to complete my PhD research.

Finally, I would like to thank my parents for their unconditional support in my life and work.

Abstract

In recent years, Online Shopping plays an important role in daily life and how to improve the online shopping experience with Machine Learning and Recommender System has been discussed by a group of researchers. As a sub-field of Machine Learning, Computer vision has achieved significant developments during the last decade. The computer vision techniques can help machine to view images and extract useful information from images like human beings. However, the existing online recommender system has mostly used the labelled information and ignored the large amount of useful information extracted from images. This thesis proposed that the extracted information from images through computer vision techniques can be used in the current online recommender system for the improved online shopping experience. To do this, I firstly tackled the problem of insufficient data in the real online shopping environment. I proposed a pairwise constraint random forest algorithm with associating transfer learning strategy. This new algorithm can make use of weakly supervised labelled data which is relatively easy to collect in the real online shopping environment to train the attribute classification model. Secondly, I developed an explainable recommender system with self-identity theory. This new recommender framework is built based on the weakly learning algorithm proposed above to analyse human behaviours by self-identity theory from information system research. Compared with previous recommender system, my work concentrates on different customer behaviours distinguished by self-identity and result in an improved online shopping experience.

In summary, there are two major contributes for this thesis. Firstly, this thesis introduces a new weakly-supervised learning approach for semantic data classification in the online shopping environment. This new algorithm can work with noise partially labelled

data to achieve better accuracy for attribute learning tasks. Secondly, by analysing the recommender system with self-identity theory, a new explainable Recommender System is proposed to improve online shopping experience. Besides, we also indicate the potential of further research in combining Computer Vision in Computer Science with online shopping experience in Information System research which can determine how Computer Vision can help to solve real-world problems.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Online Shopping Experience

The research of personalized shopping experience has attracted a considerable growth of interest in recent decades. Based on this research, there exist plenty of theories about how to optimize shopping experience (Leventhal, Mascarenhas, Kesavan, & Bernacchi, 2006; Verhoef et al., 2009) and also a lot of successful systems to improve customers' shopping experience (Freiberg, Tracy, Sanborn, & McKain, 2009). With the development of online shopping, the shopping experience of online customers has been investigated by many researchers (Bauer, Falk, & Hammerschmidt, 2006; Kuo & Chen, 2011; Lai, Shih, Chiang, & Chen, 2009; Park & Kim, 2003) from different perspectives. Among these studies an important issue of online shopping experiences lies in the difference between online and offline shopping experience (Childers, Carr, Peck, & Carson, 2002; Hernández, Jiménez, & José Martín, 2011). These researchers showed that the socioeconomic variables which are traditionally considered to be important have changed to be insignificant but security aspects tend to be more relevant. Based on those findings, online shopping websites are built to improve the shopping experience from several perspectives including quality control of website (Kuo & Chen, 2011),interface design for elderly people (Kuo, Chen, & Hsu, 2012), service quality experience (Bauer et al., 2006) etc.Previous studies have identified that product uncertainty and low retailer visibility will have negative impact on customer satisfaction and thus poor online shopping experiences (Luo, Ba, & Zhang, 2012). Researchers have endeavored to capture more information about products and other features to enhance customers' online shopping experience including utilizing big data, computer vision, and machine learning techniques recently developed. This research provides many theoretical methods on how to improve the shopping experience, but most of those methods are based on the existing techniques used in websites and do not take new features from Machine Learning and Computer Vision into consideration (Kuo & Chen, 2011; Kuo et al., 2012; Luo et al., 2012). Therefore this kinds of new features which could provide positive effects on the online shopping experience still need to be analyzed.

1.2 Online Recommendation System

The online recommendation systems for improved customer online shopping experience have gained popularity because the past transaction data could be used to predict customers' purchasing choices (Witten & Frank, 2005). At the same time, there are many successful solutions for online customer recommendation systems (Andersen et al., 2008; Robillard, Walker, & Zimmermann, 2010). For example, Amazon had increased nearly 30% of its sales by developing the online recommendation system from customer browsing history . At the same time, the online recommendation system also helps Amazon to control the security and price of the selling item by analyzing the big data provided by customers and products (Linden, Smith, & York, 2003; Rijmenam, 2016).However, most existing online recommendation systems are developed from readable text (Banko, Cafarella, Soderland, Broadhead, & Etzioni, 2007; Jacobs, 2014; Schmitz, Bart, Soderland, Etzioni, et al., 2012), leaving many new types of data such as image and multimedia data unused. Multimedia data and image data provides much richer information than readable texts. How to use those multimedia data is still a question for the Online Shopping Environment.

1.3 Visual Attribute from Images

As living in an information explosion era, both quantity and variety of data are growing significantly fast. How to effectively use this information could be a big problem. Actually, there is already some research on mine massive datasets (Witten & Frank, 2005)and this work get very successful solution (Chu et al., 2007; Thusoo et al., 2009). However, the problem of those solutions is that the focus of the research is on the technical side which does not highly relate to the marketing theories and also do not directly connect with marketing usage. Moreover, the data collected are mostly from readable text (Banko et al., 2007; Jacobs, 2014; Schmitz et al., 2012), so how to get useful information from multimedia resources like images is another problem.

In previous design of online shopping websites, the focus of images is only on how to present it clearly, as mentioned in some studies (Fiore, Kim, & Lee, 2005), the Image Interactivity technology which enables the creating and manipulation of product images has a positive impact on online shopping experience, so how to use the information from images could be a interesting research. In recent years, thanks to the techniques from machine learning, Computer Vision has witnessed some big breakthroughs. And as one of major focus in Computer Vision, the recognition of image categories has also shown some substantial improvements (Fei-Fei & Perona, 2005; Lazebnik, Schmid, & Ponce, 2006; Torralba & Oliva, 2003). The aim of this recognition is known to just classify different objects, such as a shirt, shoe or hat, some business solutions had already used this method to do image mining. However, the problem of this kind of recognition is that it usually ignores some appearance of object such as the color and texture. In order to solve this problem, some new models were introduced to learn visual attributes (Ferrari & Zisserman, 2008). By using this method, some human understandable properties could be found from images, and if I put those properties as labels attached to images, then I can group images by a combination of labels (Farhadi, Endres, Hoiem, & Forsyth, 2009; Kovashka & Grauman, 2013). For example,I can describe a shirt with black and white stripes or a white shirt with red circles on it from learning from the image. By using those methods, I could extract some features from images by analyzing personal shopping history.

1.4 Research Questions

Machine Learning and Computer vision has achieved great developments in last decade, especially the feature categorizing and detection. How to exploit the new techniques in this research area is rarely discussed in the information systems field. Theoretically, the machine learning and computer vision techniques can be applied into many areas and transform the traditional process of doing things in computer aided artificial intelligence. For example, education, health care, retailing, transportation and home development. However, application side of the machine learning and computer vision is still in the early stage. I are trying to answer one particular question for the application of machine learning and computer vision in the online environment, i.e. can I have more personalized recommendations with computer vision techniques, especially in the online environment where recommendation system have been developed for years. Specifically, our research question are:

1. is there existing research gaps between Online Shopping Experience and Machine Learning?
   Online shopping experience and Machine Learning are two popular research topics in Computer Science and Information System. Since both of them contain a lot of sub-field, is there a suitable way to combine those researches and the research gaps between them is the question I want to answer in this thesis.

2. How can I extract meaningful information from pictures from the online environment effectively?
   Most Machine Learning and Computer Vision methods need high quality data, but the currant online shopping environment cannot provide this information. I need to build more robust methods to make use of this imperfect/noise data. This thesis provides an algorithm to handle the imperfect data in online shopping environment.

3. How to explain the current online recommendation system with information theories and how those theories affect recommendation system and shopping behavior?
   The previous research in Computer Science is mostly accuracy or dataset focused which explain their model in mathematical level. While, in Information Science research, the recommendation system is usually regarded as black box that the theory behind the system is not analyzed. This thesis proposes a framework of integrating theory from Information Science to a semantic explainable Recommendation System.

4. How can the enhanced online shopping recommendation system improve shopping experience?
   Most recommendation system focus on accuracy calculated from a large amount of user. Since I introduce information science theories

to the recommendation model, I want to prove how recommendation system improve online shopping experience in theoretical level. In this thesis, I have used the Self Identity theory from information system field to evaluate and explain how the online shopping experience has been improved with the enhanced online recommendation system proposed above.

1.5 Improving Online Shopping Experience with Computer Vision

Computer vision and pattern recognition has achieved significant developments in the last decade, especially the feature categorizing and detection (Fei-Fei & Perona, 2005; Lazebnik et al., 2006; Torralba & Oliva, 2003). How to exploit the new techniques in this research area is rarely discussed in the information systems field. This project aims to explore the opportunities from the most recent developments from the computer vision area so that the online shopping experience could be improved. I discussed the possibility of extracting meaningful information from images and applying this to the online recommendation system to improve online customer shopping experience. Implications for both researchers and practitioners are discussed. The contribution of this project is twofold, firstly, I have summarized the state-of-the-art of computer vision developments in the online shopping recommendation system, especially in the fashion industry; secondly, I have provided some potential research gaps for how the computer vision method could be used in the information systems field.

In this project, I considered the problem from three different aspects: extracting semantic attributes from images (Bossard et al., 2013), enriching recommendation system with image features (H. Chen, Gallagher, & Girod, 2012a; Di, Wah, Bhardwaj, Piramuthu, & Sundaresan, 2013a; Jagadeesh, Piramuthu, Bhardwaj, Di, & Sundaresan, 2014) and image analyzing with humans in the loop (Branson et al., 2010; Mensink, Verbeek, & Csurka, 2011). I reviewed previous work in the CV area and discussed why the previous work cannot be directly applied to the current online recommendation systems. From this work I found out that, on the one hand, to apply the computer vision methods into online recommendation systems, is essential to gain insights and knowledge from the customers' perspective. More research shall focus on testing and investigating customer feedback on the current online recommendation systems through computer vision methods. On the other hand, most of the existing work relies on a large well-labeled dataset which is hard and expensive to collect from the Information System. There is a call for a method to deal with the noise and partially labeled data to obtain the desired results.

1.5.1 Improving Attribute Classification with Imperfect Dataset

Semantic attributes extracted from images could help to improve many interesting applications in information systems. However, as mentioned above, the learning of such attributes requires a large well-labeled dataset which is usually difficult and expensive to collect and sometimes requires human domain experts to annotate. Partially labeled data, on the contrary, are relatively easy to obtain from social media websites or annotated by less experienced people. A partially labeled dataset usually contains a lot of noisy data which is challenging for previous methods to process. In this project, I proposed a semi-supervised Random Forest algorithm that can handle a small well-labeled attribute dataset and large scale pairwise data at the same time for classifying grouped attributes. Results on two typical attribute datasets show that the proposed method outperforms the state-of-the-art attribute learner.

There are some previous works in this field (Bossard et al., 2013; Di, Wah, Bhardwaj, Piramuthu, & Sundaresan, 2013b; Hu, Yi, & Davis, 2015b). However, most of this research relied on well-labeled datasets which is hard to collect in real Information Systems. In order to solve this problem, one possible solution is transfer learning (Pan & Yang, 2010) or domain adaptation (Ben-David et al., 2010) which converts well trained models from a previous well-labeled dataset to the new dataset with limited labeled data. However, the limitation of both methods are that they require the well-labeled and new dataset to be related. This requirement is easy to satisfy when only category level information is considered. When applying these methods into our project, it causes problems because I mainly concentrate on attribute level categorization in Information Systems which are mostly dataset specified (e.g. dressing style in clothes dataset or sociality in animal dataset). I consider another possible solution which is to make use of unlabeled and partially labeled data that can be easily obtained from the Internet or quickly annotated by humans. Utilizing both well-labeled and unlabeled/partially labeled data is a typical semi-supervised learning scenario. However, partially labeled data usually contains a lot of noisy information, especially for abstract attributes such as the style of clothing or the design of fashion items. In our project I mainly concentrate on attribute level categorization in Information Systems which are mostly dataset specified (e.g. dressing style in clothes dataset or sociality in animal dataset).

In this work, I take the noise pairwise labeled data into the training set and use ensemble methods to make it robust to noisy data. By applying pairwise node splitting strategy to normal splitting methods and evaluating the OOB (Breiman, 2001) error when constructing trees, I generate the Pairwise Constraints Random Forest(PCRF) methods which performs better than other methods with noise dataset.

1.5.2 Self-identity Theory

Theory of planned behavior is a theory which explains how beliefs impact human behaviors. This theory shows that an individual's behavior is influenced by personal attitude, subjective norms and perceived behavioral control respectively. Among those subjective norms are the perceived social pressure to engage or not to engage in a behavior (Charng, Piliavin, & Callero, 1988). It's usually used to analyze social influence and how it impacts human behaviors (Ajzen, 1991). However, subjective norm mainly focused on social influence, when customer has a chance to experience product, and all the choices are voluntary. Self-Identity has a more significant influence on it and SN has little effect (Hogg & Terry, 2000). So this project tries to explain what is self-identity and how it determines online shopping behaviors.

Self-identity is a kind of internally generated role-expectation (Thoits & Virshup, 1997), it discusses a question "who am I in my own eyes?". This theory has a significant direct effect on the acceptance in voluntary and experienced situations, so it captures different aspects of social influences which are not captured by Subjective Norm (Sparks, 2000). Previous studies have found Self-Identity influence the individual behavior in many aspects (Granberg & Holmberg, 1990; Sparks, 2000). Compared with other social influence theories, Self-Identity has two distinct characteristics. Firstly, Self-identify captures the voluntary aspect of social influence (Thoits & Virshup, 1997) which means it initially compares other experience and transforms them into my own self-expectation, and then analyze behaviors based on the self-expectation.

Secondly, the effect of Self-Identity, unlike Subjective Norm, does not reduce with repeated experience of performing the relevant behavior (Terry, Hogg, & White, 1999). As a result, I use self-identity to describe customer shopping behaviors in the Online Recommendation System. In this project, I use Deep Learning to collect attribute level product features from visual images. Based on those features I then use self-identity to predict online customer behaviors and build a better Online Recommendation System.

1.6 Aims and Objectives

The previous research of online shopping (G.-G. Lee & Lin, 2005) shows the dimensions of web site design, reliability, responsiveness, and trust, affect overall service quality and customer satisfaction. This thesis mainly explores how online shopping experience could be improved from the web site design especially online recommendation system perspectives and explores what new type of technologies from Computer Vision could be used to improve the website design. Meanwhile, I also try to discuss how machine learning methods can be used in real online shopping environment. Those research solve the four research question proposed above and in-depth discuss the implications of computer vision technique in information systems area.

1.7 Contribution

In conclusion, by reviewing the recent research in Machine Learning and Online Shopping Environment, the research gap has been summarized in this thesis. Based on the literature review, this thesis introduces two major contributions to the current Online Shopping Experience and Recommendation System research as following.

1. This thesis introduces a weakly supervised pairwise learning algorithm to solve the problem of insufficient data in the real online shopping environment. This novel algorithm can effectively handle noise partially labelled data and generate a reasonable attribute classification model. Besides, I also propose a transfer learning strategy which can efficiently make use of noise pairwise data. In the experiment, the new algorithm outperforms the previous algorithm.

2. Based on previous findings and studies, this thesis proposes an explainable recommendation system with Self Identity theory. The explainable recommendation system combines the classification results from visual product images and the self-identity analysis from online shopping customers for a better understanding of online shopping behaviours. With the help of Self Identity theory, the enhanced recommendation framework for a better online shopping experience was developed. Compared with previous recommender system, the new proposed recommender system does not focus on statistical evaluations such as accuracy, but particularly concentrates on customer behaviour toward a better online shopping experience.

1.8 Structure of Thesis

This thesis include seven chapters which introduce the research questions, reviews on previous research, design of the experiment and the conclusions. The detail of each chapter is described as following:

Chapter 1 introduction I address both online shopping experience and machine learning especially computer vision are popular research topics in Information System and Computer Science. I also describe the potential to combine those two research fields and the problems I need to solve as our research target. In addition, I generally introduce Online Recommendation System and Self identity theory and the experiment I have done to answer the research questions.

Chapter 2 literature review I firstly review Online Shopping Experience and Machine Learning researches to introduce the previous works on how to optimize the online shopping experience by knowledge discovery in online shopping communities, and conclude the research gap and insufficient research fields in Online Shopping Experience and Computer Vision. Based on the review, I conclude our research mainly focus on two different sub-fields of online shopping experience and computer vision. The first one is Computer Vision with imperfect data from online shopping environment, and the second one is how Self Identity explain and enhance Recommendation System.

Chapter 3 methodology As a cross-field research between Computer Science and Information System, I introduce the possible research methodology of our experiments and discuss which is the best methods for our projects. Following, I describe the detail of our research protocol and the associate evaluation metrics in each experiment. In addition, I also discuss the choices of dataset selection for our project.

Chapter 4 weakly pairwise attribute learning algorithm Based on the previous works in Chapter 2, it is realized that the insufficient well-labeled data could be a bottleneck for applying CV to online shopping environment. To solve this problem, I developed a new algorithm which takes both well-labeled and partially labeled data as training set. Meanwhile, considering there are lots of noise information in current Information Systems, I also enhanced the transfer learning to be robust for the noise data.

Chapter 5 recommendation with self identity I have applied self-identity theory to fill the research gap between Information System and Computer Vision. In this experiment, I firstly extend the current online recommendation system with attribute level information for visual image data understanding. Then with personal shopping histories and related visual features extracted above, I use the self-identity theory to describe and separate online shopping behaviors into different groups. Those different behavior groups are defined by unconscious belief during shopping. Finally, I propose a new recommendation strategy considering the Self Identity of customer as a important factor.

Chapter 6 conclusion and further work Based on the experiment results, I explain how I solve the research questions proposed in this thesis and how the result prove our algorithms and solutions. I also discuss the potential to combine Computer Vision and Online Shopping Experience researches, and explain the framework I created in this thesis to help for filling research gaps. Meanwhile, I conclude all the contributes of the thesis and describe the limitations of our work. Then I describe two further works which can be regarded as the new research trend in this field. One research trend is Online Shopping with human-in-the-loop which analyzing online shopping customers with real time feedbacks. The other is a industrial project, the Roommate Matching System. Roommate matching system is a project cooperate with Road 51 company.

Online Property Renting is an subfield of online shopping that contains special factors and requirements. In this project, I plan to build a roommate ratting system with complex personal and social factors to support the original Online Property Renting website. With the system I can help tenant to find the best properties and roommates.

Chapter 2

Literature Review

This thesis concentrates on two research fields, Online Shopping Experience and the related Machine Learning research. In this chapter, we firstly introduce the definition and recent research on Online Shopping Experience, and then we indicate the important factors which can be improved by Machine Learning. In the second part of this chapter, we review and discuss what kind of Machine Learning algorithms can be used to improve the Online Shopping Experience and address the research gap between Online Shopping Experience and Machine Learning.

2.1 Online Shopping Experience

2.1.1 Introduction

The research on personalized shopping experience has attracted a considerable growth of interest during recent decades. Theories on how to optimize the shopping experience (Leventhal et al., 2006; Verhoef et al., 2009) and build some successful systems to improve customers' shopping experience (Freiberg et al., 2009). In recent years, with the burgeoning growth of various e-commerce websites, the Online Shopping Experience (OSE) has been discussed by a group of researchers (Hashim, Ghani, & Said, 2009; Hernández, Jiménez, & Martín, 2011; G.-G. Lee & Lin, 2005). In this review, we will discuss what is OSE in general and review how different researchers analyze the Online Shopping experience from different perspectives. We firstly reviewed the traditional definition of customer experience, followed by the comparison between online and offline customer experience, and lastly define the scope of our study on the online shopping experience.

2.1.2 Offline and Online Shopping Environment

Before the existence of the online shopping environment, offline shopping experiences had already been investigated by many researchers (Arnold, Reynolds, Ponder, & Lueg, 2005; Jones, 1999) from various perspectives including customer behavior and purchase intention (Jones, 1999; Quan & Wang, 2004). To trace the origins of online shopping experience, it is necessary to evaluate the differences between online and offline shopping environments, meanwhile, consider whether the existing offline shopping experience could be moved or transferred to the online environment. The key differences between Offline and Online shopping are shown in Table 2.1.

Table 2.1: Different factors between Online and Offline Shopping.

	Offline	Online
Personal Contact	Rich	Poor
Product Information	Varies	Limit
Product Options	Limit but focused	Unlimited
Flexibility	Centralized Time, ruled by Company	Centralized Time, ruled by Company

Both the online and offline shopping experience can be viewed as consequences of many interacting or independent factors. The early researchers analyze the difference between online and offline shopping experience from the different shopping contexts and drew several key conclusions (Frow & Payne, 2007; Grewal, Levy, & Kumar, 2009; Levin, Levin, & Weller, 2005). The first key difference could be viewed from the information theory perspective which is how much information the seller could get from the contacting customer. Traditional shopping usually happens in the face-to-face environment where sellers have the opportunity to find out more information about customers (Frow & Payne, 2007). This information is usually sufficient for analyzing the customers experience towards different products and providing personal suggestions. On the contrary, online shopping is fulfilled through the Internet, the communications between customers and sellers are limited. Because of this limitation, the shopping experience is hard to be customized in the traditional way. Counter-wise, customers could get as much detailed product information in the offline shopping environment but could only find limited text descriptions or pictures of products (Fatma, 2014; S. A. Lee & Jeong, 2014) in the online shopping scenario, which is the main cause of the higher perceived risk for e-customers. The third difference is the range of product options. According to the limited space of the offline sellers, the products that are available for selling are usually limited but well picked by companies. On the contrary, in online shopping, customers face a large number of products and it is usually hard to distinguish which one is the best (Grewal et al., 2009). The last difference is the flexibility of time and space. Online customers can purchase via Internet at anytime and anywhere they prefer. In the offline shopping environment, the time and location are restricted by the retailer (Frow & Payne, 2007; S. A. Lee & Jeong, 2014).

In summary, when shopping happened in the Internet context, many key variables were different and many researchers believe it is hard to directly use offline shopping strategies to evaluate the online shopping experience. Thus, it is necessary to develop new models to define, analyze and evaluate the Online Shopping Experience (Frow & Payne, 2007; Grewal et al., 2009; S. A. Lee & Jeong, 2014).

2.1.3 Customer Experience in the Online Shopping Environment

In recent years, there are many new researchers trying to evaluate the online shopping experience with much more detailed factors, but most of those factors can be viewed as a sub-factor from the online shopping environment and online customers' behavior. In the initial stage of online shopping, computer with Internet is usually considered as a hypermedia-mediated environment between customers and retailers (Hoffman & Novak, 1996), which means the internet has unique characteristics different from the traditional marketing media. As a result, Novak, Hoffman, and Yung (2000) firstly describe the differences between online and offline shopping as happened in different commercial environment, and the behavior of online shopping customers is usually outside the firm's control. From this work, the shopping experience in the online context had already been distinguished from the offline shopping experience and the major differences are the online environment itself and the corresponding online customers' behavior.

Based on the above research, there are many different definitions of OSE. One of the most acceptable concepts defines the Online Shopping Experience as the subjective response of customers to any direct or indirect contact with an online shopping progress (Frow & Payne, 2007; Grewal et al., 2009; Häubl & Trifts, 2000; Meyer & Schwager, 2007). The subject response in this concept means the response of customer's interaction with different online organizations which include the website itself, the performance of product, packaging, pricing, advertising, customer service and so on (Rose, Hair, & Clark, 2011). From this definition, OSE tries to analyze the customers' response to the Online Shopping Environment, then we can distinguish Online Shopping Experience from other related concepts such as Customer Satisfaction and Re-Purchase Intention (Kranzbühler, Kleijnen, Morgan, & Teerling, 2018). In this case, this thesis will review the OSE from two aspects which is Online Shopping Environment (Web service) and online customers' behavior.

2.1.3.1 Web Service

To identify the key factors of the Online Shopping Experience, a large number of studies focused on the website and service quality and tried to find the relationship between the quality and online shopping experience. This work has involved a range of different variables and develop different measurement methods to evaluate the web service quality.

In early web service quality research as shown in Table 2.2, the evaluation framework includes E-Qual (Kaynama & Black, 2000), WebQual (Loiacono, Watson, & Goodhue, 2002) or .comQ (Wolfinbarger & Gilly, 2002), treating the online shopping as a whole and unseparated progress with basic requirements as variables (e.g. Navigation, Accessibility etc). Meanwhile, the online shopping customers are considered to be utilitarian to the quality of service. Those kinds of early stage quality research are clearly not sufficient to describe how Web quality influence the online shopping experience. Therefore, E-S-QUAL (Parasuraman, Zeithaml, & Malhotra, 2005) and eTransQual (Bauer et al., 2006) is introduced Those evaluation frameworks not only consider the utilitarian reflection, but also investigate how hedonic emotional motives influence online shopping customers. For example, the customer experience in eTransQual is split into four phases, Information Phase, Agreement Phase, Fulfillment Phase, After-Sales Phase. Those four phases reflect four steps of customers‘ online shopping. By this splitting, the complex variables are logically sequenced, and the result is contributed to by the four phases. The way we regard online shopping behavior as a transaction makes it easy to introduce more factors for analyzing, as a result they firstly found enjoyment as a hedonic factor also important in Online Shopping Experience. Begin with this research, both customers' behavior and online shopping environment are considered, and according to the definition, the result of this work could be seen as how Web Service influences OSE. In recent years, the new frameworks usually control more variables or combine with different theories. For example, H. Li, Aham-Anyanwu, Tevrizci, and Luo (2015) combines eTailQ with Value Perception theories and Xu, Benbasat, and Cenfetelli (2013) take a lot of information quality factors into the model. However all of them contain the relationship between Customer behavior and the Shopping environment, so they could see how the Web service contribute to OSE analyzing.

From Table 2.2 it is clear to find different work usually concentrates on different variables to evaluate service quality and how this influences OSE. Among all the works, the technology part (Website design and reliability), Customer Support and Security/Privacy is always considered as an important variable, contributing to Service Quality and then resulting better OSE. Those can be seen as the basic needs of Online Shopping (Sohn & Tadisina, 2008; Xu et al., 2013). In addition, different researchers usually concentrate on different aspects of Web services, and introduce additional variables to evaluate the importance of them. Loiacono et al. (2002) firstly consider the quality of information provided by Shopping Websites and find it is highly related to Shopping Experience. Further, Xu et al. (2013) split the quality of information into different scopes (Completeness, Accuracy, Format and Currency) and figure out the importance of information quality, especially the Completeness and Accuracy. Those works find the importance of Information Quality and believe it is highly related to Service Quality and Customer Experience. However, those researchers only evaluate the basic type of information provided by Website. With the development of Online Shopping websites, there are varieties of information provided by websites, such as recommendations from big data or complex multimedia resources. So, there os some work to define and evaluate the necessity of different information provided by websites, and this work is regarded as Design Quality Evaluation. Based on the research (Loiacono et al., 2002; Xu et al., 2013), better Design Quality leads to better Web Service, and finally results in a better customer shopping experience.

Table 2.2: Service Quality Analyzing in OSE

Authors	Method	Variables									
		Enjoyment	Information Quality	Personalize	Value Perception	Website Design	Reliability	Customer Support	Security /Privacy	Fulfillment	Efficiency
Kaynama and Black (2000)	E-QUAL	X	X	✓	X	✓	X	X	X	X	✓
Loiacono et al. (2002)	WebQual	X	✓	X	X	✓	X	X	X	X	X
Wolfinbarger and Gilly (2002)	.comQ	X	X	X	X	✓	✓	✓	✓	X	X
Wolfinbarger and Gilly (2003)	eTailQ	X	X	X	X	✓	✓	✓	✓	✓	X
Parasuraman et al. (2005)	E-S-QUAL/E-RecS-QUAL	X	X	X	X	✓	✓	✓	✓	✓	✓
G.-G. Lee and Lin (2005)	SERVQUAL	X	X	X	X	✓	✓	X	✓	X	X
Bauer et al. (2006)	eTransQual	✓	X	X	X	✓	✓	✓	✓	✓	✓
Sohn and Tadisina (2008)	SERVQUAL-(extened)	X	X	✓	X	✓	✓	X	✓	X	✓
Xu et al. (2013)	3Q	X	✓	X	X	✓	✓	X	✓	X	X
H. Li et al. (2015)	eTailQ with Value Perception	X	X	X	✓	✓	✓	✓	✓	X	X

2.1.3.2 Web Design

The website and service design (also known as a part of Human Computer Interaction) are another major factor for online shopping experience (Dix, 2009; Zviran, Glezer, & Avni, 2006). We review the most used design characteristics in Table 2.3. Compared with Table 2.2 and 2.3, we may find a lot of common factors, but they are not exactly the same. For example, Efficiency in Web Design quality only refers to the efficiency of the interface, but in Web Service it also refers to the efficiency of service. This relationship means Web Design quality generally can be regarded as a factor in Web Service quality. With the development of UI design technology, we usually split Web Design quality into lots of sub-factors and analyze how those variables influence Web Design quality. According to research, better design of websites which include rich information content and ease of accessibility usually result in better online shopping experience (Ranganathan & Ganapathy, 2002). The new technologies introduced to online shopping website usually have a positive impact on the shopping experience (Ha & Stoel, 2009). In recent years, a large amount of multimedia resources is used in online shopping websites and the additional multimedia information is a necessary supplement for product description (Aghekyan-Simonian, Forsythe, Kwon, & Chattaraman, 2012; Aljukhadar, Senecal, & Ouellette, 2010). However, most of the research focused on evaluating the existing design, what kind of new technology can improve Web Design quality still needs to be analyzed and then used to guide the design of new algorithms.

2.1.3.3 Online Customer Behavior

Another group of works focusses on the online customer behavior, which includes the personal cogitative factors and the affective factors of the customer. Compared with Web Quality evaluation, those works are trying to analyze online shopping in personal feelings and behavior aspects, and they also result in OSE.

The theory of perceived risk has been discussed for many years and successfully describes lots of behaviors in different fields (Moorman, Zaltman, & Deshpande, 1992; Morgan & Hunt, 1994; Rousseau, Sitkin, Burt, & Camerer, 1998). There are also many researchers trying to use it to identify OSE. From studies (Kimery & McCord, 2002; Liao & Cheung, 2001; McKnight, Choudhury, & Kacmar, 2002; Schoder & Yin, 2000)Perceived risk shows a signifi-

Table 2.3: Web Design Analyzing in OSE

Author	Usability	Ease-of-use	Layout	Technical	Text	graphics	Visual	Speed	adequacy	navigation	security	reliability	accuracy	communication	Personal-contact	Customer-support	Currency
Lin et al. (2011)	✓	X	✓	✓	X	X	X	✓	X	X	X	X	X	X	✓	X	X
Kaya and Kahraman (2011)	X	✓	X	X	X	X	X	X	X	X	✓	✓	✓	X	X	✓	X
Yu et al. (2011)	X	X	X	X	X	X	✓	✓	X	X	✓	X	✓	X	X	X	X
Hu and Liao (2011)	X	✓	X	X	X	X	X	X	✓	✓	X	✓	X	✓	✓	X	X
Chiou et al. (2011)	X	✓	X	X	X	X	X	X	X	✓	X	✓	X	X	X	✓	X
Lin (2010)	X	X	X	X	X	X	X	X	X	✓	✓	✓	✓	X	X	X	✓
Éthier et al. (2008)	✓	X	✓	X	✓	✓	✓	X	X	✓	X	X	X	X	X	X	X
Lin (2007)	X	X	X	X	X	X	X	X	X	X	✓	X	✓	X	X	X	X
Zviran 2006 Zviran et al. (2006)	✓	✓	✓	X	X	X	X	X	X	X	✓	X	✓	X	X	X	X

cant negative impact on the online shopping experience, and the risk mainly comes from the concern of system security (Miyazaki & Fernandez, 2001), privacy infringement (Doolin, Dillons, Thompson, & Corner, 2007) and the most important, product risk (Bhatnagar, Misra, & Rao, 2000). Based on this work, further studies extended the perceived risk theory with more online shopping attributes. For example, Perceived Ease of Use (Perea y Monsuwé, Dellaert, & De Ruyter, 2004) which evaluates what website factors will lead to ease of use, Perceived Control (Chang, 2008) which refers to consumers' feelings about how they have control over their own access and search of the content in online shopping, Perceived Benefits (Sarkar, 2011) which describes the feelings required to get reward and benefit though online shopping, and Perceived Enjoyment (Suki & Suki, 2007) which refers to the hopeful pleasure felt by online customers. From this work, we may find when applying perceived risk to OSE, they usually control similar factors as Web Quality evaluation does, but constructed with different theories, which means both of them can be viewed as a part of OSE.

2.1.4 Personal Online Shopping Experience

With the development of computer technology, online shopping websites tried to separate customers into different groups to provide personalized shopping experience. Online shopping behavior happens between various customers and predefined static web pages. So to personalize the shopping experience is a good way to satisfy different customer characteristics (G.-G. Lee & Lin, 2005). One possible way to analyze customer characteristics is to separate the customer by different shopping orientations (Brown, Pope, & Voges, 2003), for example, the easy use of websites usually has a positive impact on price oriented or time conscious customers. On the contrary, experiential oriented and brand conscious customers prefer to use a more complex website which usually contains more product information. Another way to differ customers is through sociological variables (e.g. Education Level, Gender, Age) (Hashim et al., 2009; Hernández et al., 2011; Verhoef et al., 2009). However, different work gives different results, some claim the influence of sociological in the online shopping experience is very limited (Hernández et al., 2011), but others believe sociology factors play an important role in the shopping experience (Hashim et al., 2009; Verhoef et al., 2009). Comparing their results, the differences may come from the different products and different websites they investigate, generally, sociological factors have different influences on products from different categories, which makes the personalized shopping hard to analyze with limited questionnaire data.

2.2 Online Shopping Experience with Machine Learning

In the last two decades, with requirement to abstract information from large amounts of data, Machine Learning as a sub-field of Knowledge Discovery has gained a lot of attention from both academic and industrial area. Machine learning is defined as "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E. (*Machine Learning /*, 1997)". The development of machine learning in recently years has resulted in successful project in different research fields, such as Bioinformatics (Min, Lee, & Yoon, 2017), Information retrieval (Manning & ChristopherD, 2008), natural language processing (C. D. Manning, 1999) and so on. With the development of Internet and Online Shopping website, the size and complexity of data is increasing dramatically (e.g. ImageNet (Deng, Dong, Socher, & Li, 2009) and contains 14 million complex images collected from social media and Amazon (R. He & McAuley, 2016) provides more than 100 million user-item relationship data). Many traditional machine learning methods(also known as shallow learning) cannot deal with such amount of data. In this case, Deep Learning is introduced to handle the large and complex data (Lecun, Bengio, & Hinton, 2015). Compared to shallow learning methods, deep learning usually contains much more processing layers and parameters which can automatically extract features rather than use handcrafted features that is better to process complex data types (Lecun et al., 2015). Those features of deep learning algorithm make it possible to be used in Online Shopping Environment.

The current research trends of applying Machine Learning Methods to Online Shopping Environment are in three aspects:

1. Online Recommendation System with multimedia and personal information.

The original Online Recommendation System is based on big data analyzing algorithms such as collaborative filtering (Sarkar, 2011) and content-based filtering (Ricci, Rokach, & Shapira, 2015). In this stage, only the probability between customer and product is considered and the accuracy is highly reliant on the amount of data. As discussed in Chapter 2.1, there are lots of information that could be extracted from online shopping environments, but the original Online Recommendation System does not include this information because of the bottleneck of computation ability. With the development of Machine learning, analyzing more complex features such as personal information (e.g. gender, age, nationality) and product details (e.g. price, brand) becomes possible (Ricci et al., 2015). However, this additional information is mostly text data. Machine Learning methods are still hard to handle binary data such as images and videos. In recent years Computer Vision with deep learning makes binary data analyzing possible, as a result, there is a lot of works been done to include Computer Vision methods to Recommendation System (Hu et al., 2015a; Jagadeesh et al., 2014; Yue, Wang, El-Arini, & Guestrin, 2014). The new Computer Vision enhanced Recommendation Systems take a lot of personal and product information into models, which makes the model too complex and hard to evaluate. As a result, in information system perspective, how to evaluate those complex models becomes a problem.

2. Machine Learning with imperfect data.
Most Machine Learning methods including shadow learning and deep learning are built based on clean and sufficient training data. However, in online shopping context most data are imperfect and noisy, which means most Machine Learning methods cannot be directly applied to the Online Shopping Environment (Qiu, Wu, Ding, Xu, & Feng, 2016). In order to solve this problem, transfer learning (Pan & Yang, 2010) and domain adaptation (Ben-David et al., 2010) are introduced. Those methods convert well trained models from a general clean and sufficient dataset to the new dataset with noisy data in the online shopping environment. However, for complex features in the online shopping environment, it is hard and expensive to build a well-labeled datasets for transfer learning and domain adaptation. A more robust method is needed in those scenarios.

3. Analyzing customers' comments and feedback with NLP.
   NLP is short for Natural Language Processing which is a machine learning based method to understand the sentiment meaning of a phase or a paragraph (C. D. Manning, 1999). Online Shopping Environment provides a large amount of text based data, some of the data(e.g. price, brand) is easy to category and analyze but there are still some complex product descriptions and customer comments. NLP can be used to automate analyzing those kind of complex text data (Jackson & Moulinier, 2007) and generating category or numerical labels. The results provided by NLP can be used as an additional feature in Online Recommendation Systems and online shopping companies can get easy to understand feedback.

Those three research aspects attract a lot of researchers from both Machine Learning and Information System fields. In this Thesis, we will focus on the Online Recommendation System and handle the imperfect training data.

2.2.1 Machine Learning Methods

The architecture and training methods of Machine Learning is a huge topic, many concepts and models could be introduced under this topic. To make the structure more clear, we firstly introduce the different methods we used and experienced in this Thesis. Then we discuss what kind of training methods can be used to help the Online Shopping Experience, and the research gap between the existing Machine Learning methods and Online shopping environment.

In the Online Shopping Environment, there are various kinds of data such as categories (e.g. gender, brand), numerical data (e.g. age, price) and multimedia data (e.g. images and videos). To handle this complex data, joint using shallow learning and deep learning methods are necessary. The shallow learning methods, for example, Random Forest are efficient and robust to process category and numerical data. But, images and videos need deep learning based methods to process. In section, both shallow learning and deep learning methods are introduced.

2.2.1.1 Shallow learning methods

Before the wide use of deep learning, most Machine Learning methods used shallow architectures that only contain one or two nonlinear layers with extracted feature (Lecun et al., 2015). There are lots of successful and famous algorithms. For statistical learning there are Gaussian Mixture models(GMMs), conditional random fields(CRF), and maximum entropy (MaxEnt), for discriminate learning there are support vector machines (SVMs), logistic regression(LR), and multilayer perceptrons (MLPs). Those shallow architectures are effective solutions to solve simple problems and learn well-constructed features, e.g. Microsoft uses random forest to do Human Pose Recognition in Kinect (Shotton et al., 2013). However, when dealing with large and complex realistic data, they usually cannot fit them very well and the quality of the model is highly depends on the quality of extracted features. In ILSVRC2012 challenge (Deng, 2012), the result of SVMs with well-constructed feature is 10% worse than deep learning model with raw images (Krizhevsky, Sutskever, & Hinton, 2012). However, training a deep learning model need a large amount of data which sometimes is hard and expensive to collect. When there is no sufficient data or the data contains a lot noise, shallow learning methods may perform better. In detail, we introduce SVM and Random Forest as following:

Support Vector Machine The Support Vector Machine is a supervised learning model for classification. The target of SVM is to find a hyper-plain to separate two classes with largest margin. The loss function for SVM is $L(x_i, y_i) = max(0, 1 - y_i(w \cdot x_i - b))$ where x and y refers to the features of data and its label, w and b is the parameters and bias for the SVM model. The training progress is to minimize the loss function. SVM is easy to train, but it can hardly handle complex unstructured features especially the image data.

Random Forest Random Forest is kind of ensemble learning methods which combine a lot of weak predictors together for better accuracy. A typical Random Forest consists a lot of decision trees which integrated by bagging. This strategy makes the random forest robust for noisy data and prevent the model to be overfitting. The disadvantage of Random Forest is similar with SVM that it can easily handle structured data but hard to process unstructured data.

2.2.1.2 Deep Learning Methods

Historically, the concept of deep learning can be traced back to the artificial neural network, which usually refers to MLPs with many hidden layers also called deep neural network (DNN) (Lecun et al., 2015). The concept of perceptron was firstly introduced by Marvin and Seymour in 1969 (Minsky & Papert, 1969). During that time, perceptron is an algorithm for learning a binary classifier, which can be regarded as a function that maps its input and transfer to an output. As a linear model, perceptron cannot solve non-linear problem. To cope with more complex data, MLP or ANN was proposed. ANN is a kind of feed-forward model that mapping input data to multiple layers of nodes and give a set of outputs. ANN usually contains many parameters which is hard and expensive to train. Back propagation was then introduced (Schmidhuber, 2015) in 1988 to use gradients to enable training large and deep ANN possible. With back propagation, ANN can be used to solve many real problems. However, back propagation also has vanishing of gradient problem (the gradients decrease exponentially from layer to layer), especially when the ANN contains more than one layers. From this point, traditional ANN only contains one hidden layer, which still belongs to shallow learning models. In recent year, with the development of architectures and training methods, building a deep ANN has been possible. We introduce several deep learning models which we used in thesis.

2.2.1.3 AutoEncoder

The AutoEncoder introduced by Hinton and Salakhutdinov (2006) is a special type of DNN that consumes unlabeled data and output vectors which is a better representation of the origin data. The AutoEncoder is typically used for dimensionality reduction or generative models.

When there is more than one hidden layer, the AutoEncoder can be seen as a deep learning model. The architecture contains an input layer which represents the input data, then connected to one or more hidden layers to transform features and finally an output layer which is usually a better representation of the input data. The whole structure contains two parts, the encoder which maps $x \in X$ to $z \in F$ and the decoder which maps $z \in F$ to $x \in X'$, the X refers to the origin space and F refers to the feature space. So the objective of AutoEncoder is to minimize reconstruction errors (Bengio, 2009) which is $\mathfrak{L}(x, x') = \left\|x - x'\right\|^2$. From the objective function, it is clear that AutoEncoder is an unsupervised learning methods. The dimension of the hidden layers can be either smaller or larger than the input dimension, when smaller, the AutoEncoder will do feature compression, otherwise, it will map features to a higher-dimension. By selecting different dimensions of hidden layers and different objective function, the AutoEncoder can do various jobs such as denoising (Vincent, Larochelle, Bengio, & Manzagol, 2008), sparsing (Le, 2013) or generating. Since we use the minimize reconstruction error as the objective function, the training of AutoEncoder is still based on the back propagation. As a result, the vanishing gradient problems with the use of back propagation to train networks with many hidden layers still exists. To solve this problem,Hinton, Osindero, and Teh (2006) produced a layer by layer pre-trained technique based on a restricted Boltzmann machine to pre-train the weights and then fine turn them by back propagation.

The AutoEncoder can help to reduce noise data in the online shopping context. The images from different website are usually in different quality and may contain mistakes, AutoEncoder is a potential way to cope with the quality and noisy problems that makes other Machine Learning algorithms works in the online shopping environment.

2.2.1.4 Convolutional Neural Network

As described above, the number of parameters in a network is directed based on the number of layers and number of neural in each layer, so when the net is getting deeper and larger, there will be too many parameters to learn which is usually hard and expensive (Krizhevsky et al., 2012). The Convolutional Neural Network(CNN) solves this problem by sharing weight in convolutional layers. It is especially good for handing complex data e.g. images, and the Convolutional layers can be regarded as a learned feature extraction method which avoids building hand-craft features. To limit the number of training parameters, in 1990s CNN (LeNet-5), a ANN with additional Convolutional layers,was introduced by LeCun et al. (1995). In the Convolution layer, the output is generated by the input convolute a small kernel, and in the pooling layer, the input will be sub-sampled to lower dimension with some strategy (max polling, average polling, etc.). Compared with full-connected ANN, LeNet-5 contains two convolutional layers in which the number of parameters needed to train only equals to the size of convolutional kernels multiply the number of kernels $h * w * d$ (e.g. in LeNet-5 the parameters in (32*32*12) plus (16*16*12). In this case, CNN significantly reduces the number of parameters. As a result, LeNet-5 has a very successful result on MINST dataset. However, LeNet-5 is still hard to train on huge and complex datasets (e.g. ImageNet). The reason is the vanishing of gradient in deep structure, and the limited layers (only 2 convolutional layers) reduce its ability. In 2012, AlexNet, a more deep CNN, was produced (Krizhevsky et al., 2012) and has a much better result than shallow learning methods (SVMs etc.). On ILSVRC (Deng et al., 2009). AlexNet, as shown in Fig 5, is known as one of the most successful deep learning architectures in recent years. Compared with LeNet-5, AlexNet choose a more deep structure and made the following main changes:

- Data Augmentation: AlexNet pre-processes images with horizontal reflections, random crops and color transformation. With data augmentation, there will be more data for training which will reduce the overfitting problem.

- Dropout: Dropout can randomly break connecting when training CNN which makes the CNN try to get the right result with a subset of neural. Compared with L2/L1 regularization, the use of Dropout can significantly reduce the overfitting problem.

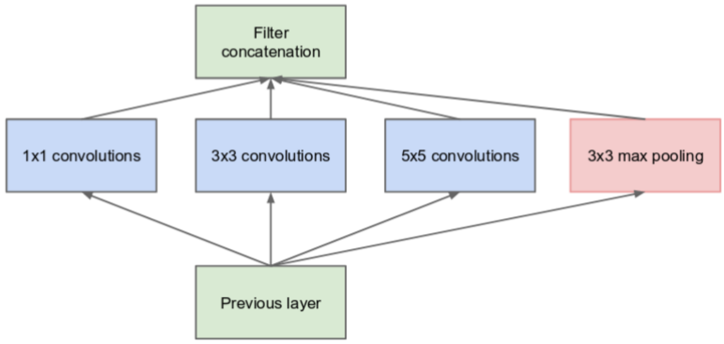
- ReLU activation function: Instead of using Sigmoid or Tanh, AlexNet use ReLU $f(x) = \max(0, x)$ as the activation function. The advantage of ReLU is two-fold, on one hand it is easy to calculate for both forward and backward, on the other hand the linear right part avoids the vanishing of gradient and the left zero part makes the network sparse, which is similar to L1 regularization.

- Local Response Normalization: Instead of normalization on all the data, AlexNet chose to normalize on recent training data.

- GPU: the ability of using GPU to do fast matrix computation is one of the most important foundations of the deep learning. Due to the length of the thesis, we will not include the hardware parts of deep learning.

After AlexNet, there are many successful deep learning architecture proposed, such as VGG (Simonyan & Zisserman, 2014), GoogLeNet (Szegedy et al., 2015) and ResNet (K. He, Zhang, Ren, & Sun, 2016). These works use different strategies and make the CNN more accurate and deep. A very successful network is ResNet which use identity mapping with batch normalization to add the origin data to output($H(x) = F(x) + x$). With this simple strategy, the problem of vanishing of gradient can be avoided, as a result the ResNet could contain unbelievable 152 layers.

Meanwhile, instead of making the CNN deeper, Inception model was also proposed to add the different size of convolution filter into the same layer,



(a) Inception module, naïve version

Figure 2.1: Basic Inception Convolution Block (Szegedy et al., 2015).

which can be regarded as extending the width of network (Szegedy, Ioffe, Vanhoucke, & Alemi, 2017). The basic Inception block is shown in figure 2.1. However, the Inception model significantly increases the complexity of the architecture, which prevents it from becoming wider and deeper. To solve this problem, ResNeXt (Xie, Girshick, Dollár, Tu, & He, 2017) was introduced, which replaced the complex inception block with a group of repeated convolution filters, and those filters can be computed by group convolution. Compared with the complex Inception model, group convolution is much quicker for the GPU to calculate. The complexity of ResNeXt can be defined and described by Cardinality, which is the number of same filters in one block. A typical ResNeXt usually comes with 32 Cardinality blocks that are much wider than Inception network.

As described above, the three typical ways to improve CNN architecture is to make CNN deeper, wilder and in higher resolution, but the bigger CNN usually means it is challenging to train and fine-tuning. Therefore, there is a trade-off between the deep, wide and resolution of a CNN network. EfficientNet (M. Tan & Le, 2019) was then introduced by experimenting and calculating the best deep, wide and resolution combinations of a CNN network and it is believed to be the state-of-the-art architecture considering the size, complexity and accuracy at the same time.

By building the CNN network, complex attribute learning becomes possible. In this case, we can extract many different middle-level features from online shopping products' images. With those features, we could build more accurate models to describe customer shopping behaviour and analyzing shopping history. Besides, theories from Information Systems or Marketing can also be easily applied to Online Shopping environment.

In our experiment for semantic attribute learning, considering the number of attributes and the size of the dataset, the slightly increased accuracy is not that important. Instead, the easy to get a pre-trained model, the hardware requirement for fine-tuning and the speed to training the model are all critical perspectives. In this case, the CNN architecture used in this thesis may be slightly changed according to different datasets, and the state-of-the-art network may not necessarily be used because our research is not accuracy focused. The details of our CNN network is described in Chapter 5.

2.2.2 Machine Learning with Imperfect Data in Online Shopping Environment

As described above, data from the Online Shopping Environment usually contains noise and errors. Semi-supervised learning is a class of machine learning techniques that makes use of unlabeled or partially labeled data for training, and those unlabeled and partially labeled data can be generated from noisy data in Online Shopping Environment. There are various data structures can be regarded as weakly or semi-supervised learning problem. For example, the visual relations in the same image can be learned with only object level label and without annotated relations (Peyre, Sivic, Laptev, & Schmid, 2017). Another weakly supervised learning task is to handle such a deep learning model with constraints incorporating different forms of weak supervision and can share those weak supervision from different datasets (Zhukov et al., 2019). In this thesis, we consider another situation that a specified partially labeled dataset as pairwise constraint data and extend the existing work to handle pairwise constraint relations.

To utilize pairwise constraints, some researchers consider it as a clustering problem and develop constrained clustering algorithms based on K-means such as COP K-means (Wagstaff, Cardie, Rogers, Schrödl, et al., 2001) and CMWK-Means (de Amorim, 2012). With a similar idea, a more efficient Semi-supervised Maximum Margin Clustering method was also introduced to handle the pairwise constraints (Hu, Wang, Yu, & Hua, 2008; Zeng & Cheung, 2012). The drawback of these algorithms is that they make a strong assumption that the pairwise constraint data are accurate. However, in practice, pairwise information is usually collected from social media sites which makes it contain a lot of noisy data. To avoid the influence of noisy data, the Constraint Propagation Random Forest which attempts to prevent the noise impact by an ensemble is formulated (X. Zhu, Loy, & Gong, 2015). Compared to the previous algorithms, it performs better when dealing with noisy data. However, the limitation of these algorithms is that they only use the pairwise constraint data to do clustering rather than classification. To build attribute classifiers, a small set of well-labeled data should be utilized along with a large amount of pairwise data.

Taking labeled data into consideration, Liu et.al. (X. Liu et al., 2013) proposed a Random Forest based approach which trains trees with both labeled and unlabeled data, and as an ensemble method it is also robust to noisy data. However, this algorithm does not take the relation among unlabeled data into consideration, but only uses them as background knowledge to find better splits. To use labeled data and pairwise information at the same time, Convex Pairwise Kernel Logistic Regression which builds a loss function with pairwise constraint information was introduced by Yan et.al. (Yan, Zhang, Yang, & Hauptmann, 2006). Similar with this work, a Margin-based approach was also proposed (Nguyen & Caruana, 2008). In this method, a regular multi-class SVM is extended with a pairwise optimization objective function. These two algorithms use both well-labeled and partially labeled data, but according to our experiments, they are not robust to noisy pairwise data and do not consider the different qualities between well-labeled data and pairwise data.

2.2.3 Recommendation System with Computer Vision

Analyzing the customers' behavior from their shopping history and using this information to make recommendations to customers so that customers' shopping experiences could be enhanced has become a trend in most e-commerce websites. (Childers et al., 2002; Park & Kim, 2003). Currently, most online shopping websites such as Amazon and ebay, make suggestions to their customer by analyzing customers' searching or shopping history. This method is successful because related items or products similar to those from their browsed history could be pushed to customers. The limitation is that all the predictions are only based on the item-to-item or user-to-item combinations (Linden et al., 2003; Poon, Maltzman, & Taylor, 2012).The algorithm of these models only considers the relations between item and user or item and item, but ignores the features of the products themselves.

The most salient features extracted from images in e-commerce websites would be used to enhance online recommendation systems and thus shopping experience. Extracted features could be those descriptive categories perceived by human beings such as color and style etc. For example, clothes on the Amazon website usually contains 5 labels: color, style of sleeve, material, brand and price, but from the pictures provided by the website we can extract more than 10 additional labels such as length, cut, pocket, collar, and material etc. (H. Chen et al., 2012a; Di et al., 2013a). Moreover, new algorithms could be built based on some public training datasets (Bossard et al., 2013; Kalantidis, Kennedy, & Li, 2013), and well trained models can automatically extract the clothing part and analyze possible labels from each type of clothes. These labels could be implemented from human perspectives and some cognitive factors could also be used to extract useful information from clothes pictures. For instance, personality type could be used to classify clothes style based on attributes extracted from images. There is thus a possibility to provide more accurate description of products from a higher cognitive and conceptual level so that customers could be provided more enriched products information at a higher conceptual and cognitive level.

The overall trend for online fashion recommendation systems enables the online shopping systems to be more personalized. There are some successful examples for the fashion recommendation systems through the combination of both text and image features. Jagadeesh et al.(Jagadeesh et al., 2014)proposed a fashion recommender by analyzing the color model from street images for item recommendation. Iwata et al Iwata, Wanatabe, and Sawada (2011)collected text and image data from fashion magazines to build a topic based recommendation system.These two works are item based which only considers the relationship between items and the item-user relationship is not considered here. With the development of social networks, personalized recommendation systems with image features are gaining popularity in recent research.Sigurbjornsson, Borkur and Van Zwol (Sigurbjörnsson & Van Zwol, 2008)proposed a personalized tag recommendation system based on a Flickr dataset. In this work, they analyzed the frequently used tags of customers to automatically recommend personalized tags for newly added photos. Another research from Yisong et al Yue et al. (2014) provided a similar personalized recommendation system by collecting customers' feedback. This type of research mostly concentrates on the customer side, and provides recommendations by finding similar customers. Meanwhile, there is also some research considering both user-to-item and item-to-item relationships at the same time (Hu et al., 2015a). In Yang et al.'s research, they built a model

Figure 2.2: Finding tops to match with given bottom and shoes with image features (Hu et al., 2015a).

with each customer's preferred fashion items and combined a set of fashion items themselves, and then made a personalized recommendation with a set of fashion item as shown in Figure 2.2.

As shown in Figure 2.2, researchers built various recommendation systems through mining the large set of data collected from computer vision methods. Meanwhile, other research trends try to include other information and product relations into the current model. Quadrana, Cremonesi, and Jannach (2018) proposed a sequenced recommender framework which takes a list of behaviours as the prediction resource. Compared with previous work, they include time-series data into current recommendation model and believe the new framework will outperform the previous one, which only works on a single behaviour. Similarly, a graph-convolution based network (Ying et al., 2018) was introduced to direct handle relation of products by CNN itself.

However, the current contribution of these new papers is mostly on the new mathematical methods or algorithms that could handle different types of datasets. These works only focus on the recommendation algorithm from the technology perspective. How customers will respond to this new type of data has not been investigated from the information systems perspective. What type of features shall be extracted? Which features are more salient in improving online customers' shopping experience still have not been explored.

2.3 Recommendation System with Self Identity Theory

In order to analyze and evaluate the Computer Vision enhanced Recommendation System we introduced Self Identity theories which helps us understand how features from computer vision works and evaluate the recommendation system results.

Early studies on online recommendation system rarely consider image as an important factor but only display the pictures clearly to achieve the best product effects (Fiore et al., 2005). The information in the picture is not fully explored mainly because the image processing techniques haven't been fully developed in early days. Alongside the development of the Image Interactivity technology which enables the creating and manipulation of product images, the potential to exploit more feature from images increase. In the beginning, researchers started focusing on sketching and modelling fashion items(H. Chen, Xu, Liu, & Zhu, 2006).Recently, due to the techniques from machine learning, Computer Vision is witnessed some big breakthroughs. One of the major breakthroughs in Computer Vision is the recognition of image categories (Fei-Fei & Perona, 2005; Lazebnik et al., 2006; Torralba & Oliva, 2003). The first improvement comes with feature representation of images, for example, at the feature level, there are kinds of features that could be extracted by different methods including SIFT (Ke & Sukthankar, 2004) ,GIFT (Oliva & Torralba, 2001), Histograms of Oriented Gradient(HOG) (Dalal & Triggs, 2005), Local Binary Pattern (LBP) (Ahonen, Hadid, & Pietikainen, 2006), Maximum Response Filters (Varma & Zisserman, 2005)).Based on these features, a well-trained model could be developed to classify different objects, such as a shirts, shoes or hats into categories. The semantic attributes provided by researcher can be used to further assist object classfications. Some business solutions had already used this method to preform image mining and achieved satisfactory results (Bossard et al., 2013). However, the problem with this kind of recognition mechanism is that it usually ignores certain type appearance of objects such as the color and texture. In order to solve this problem, some new models were introduced to learn visual attributes (Ferrari & Zisserman, 2008). By using this method, human understandable properties could be extracted from images. If we put those properties as labels attached to images, then we can group images by a combination of labels (Farhadi et al., 2009; Kovashka & Grauman, 2013)For example,we can describe a shirt in a specific style with black and white stripes or a white shirt with red round on it and classify clothes with these properties. By using those methods, we could extract some high level semantic features from images such as clothing style, patterns and textures. But these methods only work well with clear and simple image data. As a result, in the realistic online shopping environment, those methods can hardly handle the complex and noisy image resources.

In order to solve this problem, some object detection models have been developed (X. Chen et al., 2014; Yamaguchi, Kiapour, Ortiz, & Berg, 2012).These models use human pose estimation or simple object detection method to locate the interesting item in an image so that attribute learning method can be applied only to those located item. With this kind of preprocessing method, we could extract semantic attributes from images in a real online shopping environment. There is already some success research in this area. Actually, there are already some success research on this.For example, through collection of a well labelled dataset,Chen et al. (H. Chen et al., 2012a) extracted complex semantic features from clothing. Moreover, Liu et al.(S. Liu et al., 2012) collected both top and bottom clothes and identified the semantic feature relations between them, which enable them to make further suggestion on item combinations of clothes.

As shown above, applying those information collected from Computer vision method could help to improve the design of website and improve not only the description of products but also the shopping experience. However, the training of a semantic model needs a large set of well-labeled data and as the fashion items increase every year, there is also a strong demand to update the current model with new data. The previous work mostly generated their datasets with the help of crowdsourcing which is time consuming and expensive. So it is important to design attribute learning algorithms that do not require lots of well-labeled data.

2.3.1 The Definition Self Identity

In psychology, Self Identity is known as a theory that describes how social environment affects self and how self affects behaviors (Hogg, 2001; Johnson, Venus, Lanaj, Mao, & Chang, 2012; Stryker & Burke, 2000). And the behavior, in Self Identity, is organized to change the situation of self in order to bring themselves into an agreement with some identity standards (Stryker & Burke, 2000). Accordingly, Self Identity research usually contains two portions: one focuses on characteristics of a person which differ him from others (Johnson et al., 2012) and the other analyzes the link between social structures and self behaviors (Carter, 2015). The characteristics of a person contains both conscious and unconscious aspects which affected by different social or environmental structures and result in different behaviors. The analyzing of self identity is to distinguish how the conscious and unconscious characteristics may affect the behaviors (Horowitz, 2012; Nurra & Oyserman, 2018).

In previous research, Self Identity was widely used to discuss how personal and environmental differences like culture and gender affects academic and physical abilities and skills (caglar, 2009; Trautwein, Lüdtke, Marsh, & Nagy, 2009). In order to avoid unrelated factors and focus on specific targets, some researches use one portion of Self Identity theory along with environmental or social factors to build a more realistic model (Papacharissi, 2010; Van der Werff, Steg, & Keizer, 2013; Whitmarsh & O'Neill, 2010). In those research, Self Identity is a set of factors attached to the self as a standard that guides behaviors in situations. When identity is activated, a feedback loop is established where the standards defined by self identity and environmental factors will result in behaviors, on the other side, the behavior can affect the environmental factors and finally reflect to Self Identity (Stets & Burke, 2003). Therefore, Self Identity is not constant, it may changed during time and different environmental factors.

The Self Identity determined by both conscious and unconscious factors. In psychology, self-schema refers to unconscious generalization of self, and self-representation refers to the conscious belief such as gender and culture (Brewer & Gardner, 1996) which can be symbolized in words. Most previous research focus on the self-representation to build Self Identity models (Van der Werff et al., 2013; Whitmarsh & O'Neill, 2010), but self-schemas are hard to collect and analyze. Each individual has multiple self-schemas which can be activated under different situation and shift person's mind. Self-organization is the collection of self-schemas for a person. People vary in how well they realize their self-organization under different situations. The level of self-organization can be separated in to five stages: Harmonious, Mildly conflicted, Vulnerable, Disturbed, Fragmented (Horowitz, 2012). From Harmonious to Fragmented, people will less realize their self-organization and become unpredictable to their action. Therefore, the level of self-organization under a situation refers to how well the behavior can be predicted and how clear the person knows what he wants.

In conclusion, Self Identity describes how environment affects self and how self identity decides the behavior. The progress of it is not constant, it may vary from time and different situation. Self-organization is a collection of unconscious factors of Self Identity, the level of it can be used to describe how well a person known himself.

2.3.2 How Self Identity is used in Recommendation System

As reviewed above, Self Identity is rarely used in analyzing Online Shopping Experience(OSE). There may three reasons: firstly, the research on OSE mostly focus on a group of customers(under same age or culture and so on) on a specific scenario rather than analyzing one customer's behavior. Thus, group identity or social identity is more likely used. Secondly, unconscious factors in Self Identity are hard to collect and exam in normal ways, but symbolic conscious factors are more preferred to be used in OSE research (Van der Werff et al., 2013; Whitmarsh & O'Neill, 2010). Finally, the Online shopping scenario is a complex system, Self Identity theory itself is not sufficient to describe or analyze the whole system. However, to increase OSE we do not need to increase the whole system but focus on a specific aspect. In this case, we could use Self Identity to describe and predict customer behaviors under a special situation. Relative to previous OSE research, we use Self Identity theory to focus on very personalized Online Recommendation problem, and as a result, increase the online shopping experience.

Early studies on online recommendation system involve two directions: one is technical part which focus on algorithms and try to increase the predict accuracy, the other is information system part that analyzing customers' behaviors and feedbacks. The latter researches consider online recommendation system as an important factor of OSE, and the target of research concentrate on customers' feedback. The conclusion from information system research about online recommendation could be helpful to guide the development of recommendation system in computer science, but there are rarely research integrate those two parts together. Self Identity theory was never used to analyzing online recommendation system in both research directions. But considering the limitation of Self Identity theory described above, the research from CS can focus on personalize customers and collect unconscious factors(by Machine Learning). Future, with concentrating on Recommendation part instead of the whole system, Self Identity can be used to describe customer behaviors under a specific situation.

In our project, we meticulous analyzing customer product selection progress along with text and image information of products to investigate the stage of self-organization, and then obtaining the Self Identity model of each customers in each online shopping progress. The result from Self Identity, on the one hand can be used to increase current online recommendation system, on the other hand it could be used to analyzing how customer behaviors can be affected by recommendation and other online shopping factors.

2.4 Research Gaps

Based on the literature review from both Information System and Computer Science research, the research gap can be summarised, and the potential methodology to fill the gap can also be discussed.

In Computer Science research, the research gap exists in the different research target and the tight requirement for machine learning algorithms.

1. Researches taken in CS field usually focus on a specific mathematic evaluation metrics(e.g. Accuracy, F1 score), and the target of new generation algorithms is to achieve better score under those evaluation metrics. However, in Information System research, some complex research targets, such as customer behaviour, customer satisfaction, can not be described well by simple evaluation metrics used in CS research. Therefore, how to combine different research target is one important research gap between CS and IS.

2. Most Machine Learning especially Deep learning methods require a large well-labelled dataset for training. In real Online Shopping Environment, those high-quality dataset is usually different and expensive to collect. How to make existing ML algorithms working with weakly and imperfect labelled data is another research gap.

In Information System research, the research target is mostly customer-focused, and the whole progress of online shopping is usually split into different subsystems for further analysing (e.g. Recommendation, Web Service and Quality control). However, considering the research in CS, there are still some research gaps exist.

1. Those split sub-systems are usually treated as whole black boxes, and the algorithms behind them are rarely discussed. How the details of the algorithms can affect customers is not be well considered either. In this case, the influence of different CS algorithms work in the same sub-system can be hardly analysed.

2. The research of IS usually focus on the existing system, how new algorithms in CS can be applied to the current system is rarely considered. Meanwhile, the IS theories about how CS methods affect customers are not well discussed either.

2.5 Summary

In this chapter, we review the recent research in Online Shopping Experience and explain the key factors include Web Service, Web Design and Customer Behavior. Based on review, we particular discuss the importance of Recommendation System which plays an important role in the all three aspects and we also point out the limitation of previous research is that they regard Recommendation System as a black box without considering how it recommend. Meanwhile, we introduce the existing research trend about Machine Learning in Information system and basically discuss the advantage and disadvantage of different Machine Learning algorithms. The imperfect and insufficient data in real online shopping environment limit the performance of Machine Learning, especially Computer Vision. How to handle those imperfect noisy data is our first research topics. The Recommendation System is developed in both Information Science and Computer Science, but the aim of those research is different. In Computer Science, those researches are accuracy concerned which try to improve the prediction ability, while in Information System, the research is customer focused that explained how customer reflect to the Recommendation System. Thus, we introduce Self Identity and use it to explain the Recommendation System and try to improve the shopping experience with new Recommendation System.

Chapter 3

Methodology

In this chapter, we generally discuss the methodology we used to solve the problem we addressed in Chapter 1 and 2. In addition, we describe the evaluation methods for experiment and the data collection for each projects.

3.1 Introduction

As a cross field research, the methodology we could use in experiments contains both Computer Science and Information Science methods. The possible methodologies list as following:

Mathematical modeling and proof method This is a common research method in Computer Science, especially in developing new algorithms. The mathematical model will describe the input and the reflection of that input with functions, and introduce a combination of functions which will cover all interest factors (Mudaly & De Villiers, 2004). Based on the model, we could analyze and predict how algorithms working on it and how the target reflect on the algorithm with mathematical proof. The first advantage of this methodology is it abstract the important factors and hide irrelevant details from the real world (Mudaly & De Villiers, 2004). And those abstraction usually describe a common problem solving approach which is not environment and problem specified. The other advantage is those mathematical model can be mathematical proven which is very reliable and explainable (Mudaly & De Villiers, 2004). However, for complex problems with hidden factors, mathematical modeling is usually hard to abstract and define. The mathematical modeling prefer to solve common questions rather than analyzing real world problems.

Simulation The methodology of simulation refers to the approximate imitation of a process or system (Jerry, 2005). It is usually used to simulate complex models which are hard or expensive to repeat in real world. The simulation study is used in both Computer Science and Information System areas (Law, Kelton, & Kelton, 1991). The advantage of simulation is it is easy to control and collect data, meanwhile, it can hide irrelevant information from real world which let the experiment focus on target. However, to simulate the system, we usually need to collect enough information of the system and approximate all key factors (Law et al., 1991).

Quasi-Experiments Quasi-experiment is a popular research methodology used in Information Science. Compared with formal experiment, quasi-experiment just estimate the causal impact of a factor on target without strict control all other factors and the design of quasi-experiment is mostly based on empirical knowledge (Cohen & Ledford Jr, 1994). The advantage is clear that quasi-experiment is easier to set and collect results which is usually used to confirm the research trend or judgement. The disadvantage is those research include a lot randomizations and the result is hard to be clearly described and proofed.

Prototyping Prototyping is a common research method in Computer Science to test how a product works in real world or display the project to customers. In detail, the prototype only implement key functions of a system and prove the correctness of design (Houde & Hill, 1997). Compared with previous methodology, prototype is a more industrial related design methods but it can also be used to prove principles.

Survey method The survey method is mostly used in Information Science to collect the individual response on a particular topic (Nicolas, 2004). This methodology directly connect researchers with research candidate and could quickly get personalize response for analyzing that is useful for customer centered research. However, it is usually expensive and hard to collect large amount of data, and the process of survey is mostly time consuming.

Qualitative methods The qualitative research is a popular research method in Information Science. It can handle non-numerical data and conclude the meaning or concept behind those data, and those researches usually result in new definitions, characteristics or concepts (Maxwell, 2008). The qualitative method is mostly social and human centered. In experiment, it usually works with survey methods to explain the theory behind the survey.

All the six methodologies are popular used and sometimes researchers combine some of those methods together to get a better understanding to the question and establish a more meaningful model. In our research, we apply Machine Learning to online shopping environment where we could get a lot of numerical data to analyze and the results of experiment are mostly described in mathematics. In this case, during research we use simulation to approximate the online shopping environment and extract related factors, and apply mathematical modeling and proof method to establish and prove the Machine Learning algorithm. The result of our research will be evaluated by mathematical methods. The detail of our methodology approach is described from the abstract of the problem, experiment and evaluation methods to the details of design and simulation for each experiment. In addition, we also discuss the data we used in our experiment.

3.2 Overall of the Problems

In chapter 2, we reviewed the past research about online shopping experience, and mainly discussed web service,the design of online shopping website and online customer's satisfaction, especially how recommendation affect online shopping experience. Then we review the research from the Computer Vision field which may be applied to improve online shopping experience and indicate the research gap that most Computer Vision methods need well organized data which is hard to find in online shopping environment and the research in online shopping experience rarely consider how new computer vision can affect the customers' experience. In literature review, we reviewed most top conferences on Computer Vision such as Computer Vision and Pattern Recognition and ACM Multimedia Conference, especially targeting at the fashion and clothing area. The review results demonstrated that attribute learning method could be used to improve online shopping experience. We prove this by demonstrating how fashion item recommendation system could be developed with attribute learning method in computer vision.

Based on the literature review, to improve online shopping experience we could rather improving machine learning algorithms used in online shopping environment or doing theoretical analysis based on Information System theories. In this case, we split our experiments into two parts. Firstly, we try to solve the imperfect data problem in online shopping context. Then we apply self identity into the current online shopping recommendation system to describe shopping experience with self identity theory and investigate how the theory affect online shopping behavior.

Both of the problems contain lots of numerical data to analyze, and as a complex online shopping system, abstracting necessary key factors is important for our research. The survey method with qualitative analysis can not be used in our model because we collect data from public dataset rather than questionnaire and the data type is mostly numerical. Similarly, Quasi-Experiment method which focus on information science to get research trend is not suitable for our target. Prototyping is also not necessary for our experiment, as it usually concentrate on industrial projects and our research is mostly theoretical focus. In our research, we use both mathematical modeling and proof and simulation methodology to guide our design. The reason is in two fold, on the one hand, our research produce new algorithms and systems which can be modeled and described by mathematics and the behavior of our new design can also be predicted by the proof of mathematical model. On the other hand, we need abstract and simulate the complex online shopping environment to focus on the key factors we interested in. The simulation methodology can be used to do such abstraction.

Figure 3.1 describes the overall experiment structure in details. The whole experiment can be separated into three stages. The first stage is semantic attribute learning stage where the visual attributes from product images can be learned by CNN and prepared for further analysis. This stage is described in Chapter 4. The second stage is self-identity analysing, in which stage the self-identity of customers can be analysed by combining attribute information and customer-product relationships. The final stage is to analyse and build a recommender system with knowledge from previous self-identity analysis. Specifically, each input and output for each block are described as following:

Circle 1 Attribute Classifier is trained by open and public datasets.

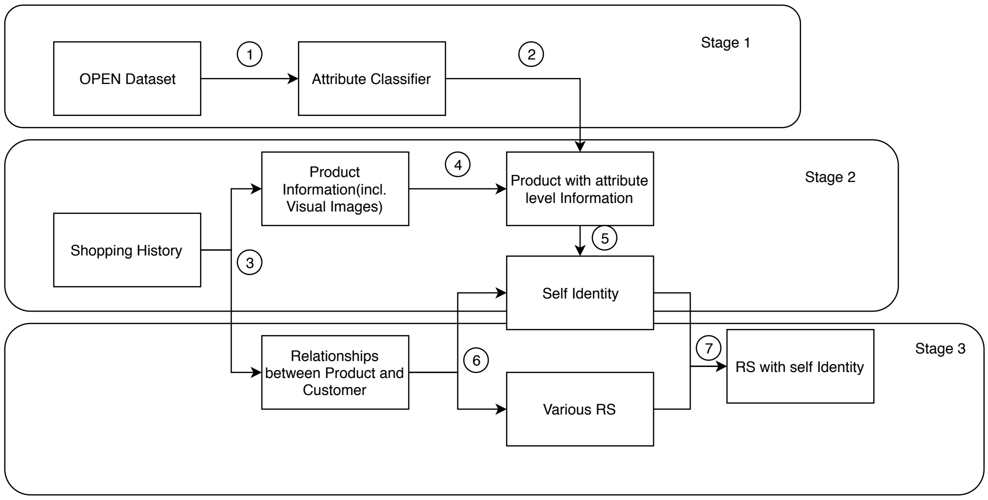


Figure 3.1: The flowchart of overall experiment design

Circle 2 The trained classifier predict visual images from product to text attribute labels.

Circle 3 The shopping history contains both product information and customer-product relationships.

Circle 4 The visual images from product information will be transferred into attribute labels.

Circle 5 Products with attribute level information along with shopping relationships will be used for self-identity analysing.

Circle 6 The product-customer relationships will be used to help self-identity analysis and build recommender systems.

Circle 7 Recommendation with self-identity theory will be finally built based on previous RS and self-identity analysis.

3.3 Experiment Method

For each experiment, we follow the research protocol as following: we firstly define and identify the problem and review relevant literature. Secondly, based on the methodology we used, we construct the experiment design and abstract the key factors in experiment. Thirdly, we use mathematical proof or simulation to introduce the new algorithm or system. Finally, we conduct the experiment with real data from online shopping environment to confirm our prediction, meanwhile, we present findings and conclusions for the target.

The define and identify problem and literature review are been generally discussed in the Chapter 1 and 2. In experiment, a more clear definition of problem will be given and the review will concentrate on related research in recent years. After that, a mathematical model or simulation of the problem will be introduced.

In second and third phase of the protocol, for mathematical modeling and proof method, the correctness of our algorithm is proven by the mathematical derivation and the experiment results from simulation. For simulation method, we abstract necessary data and key factors from the complex online shopping environment, then we use mathematical model to evaluate how our theory affect the system. And based on the results, we propose new system which is also proven by mathematical model and the experiment results confirm our design.

Next, we will implement the mathematical model based algorithm or simulation system to confirm the algorithm and system we presented. The result of our experiment will be evaluated by mathematical evaluation metric or statistical evaluation methods.

Finally, we will explain our results compared with related methods and make the conclusion for our experiment. The correctness of our claim is confirmed by both mathematical model and simulation result.

3.4 Evaluation Method

In research, we firstly build or simulate an online shopping environment with public data collected from internet. In this environment, we abstract the necessary data and how current algorithm or system works, then we analyze the weak point of the system and describe how we could improve the online shopping experience by updating the algorithm or system. Finally, we introduce our design or algorithm into the online shopping environment and compared the new one with previous system.

In Experiment, we follow the protocol of analyzing problem, building new solution and experimenting new system compared with the old one to prove our new solution can improve the online shopping experience. An important part of those protocol is to use suitable evaluation methods which can explain and compare our result with other researches. In this section, we introduce the evaluate method we used in our experiment.

3.4.1 Classification Accuracy

The Classification Accuracy is the basic evaluation metric for machine learning models. The accuracy is the fraction of correct predictions in our model as shown in equation 3.1.

$$Accuracy = \frac{|P(X) = y|}{|X|} \tag{3.1}$$

3.4.2 K-fold Cross Validation

The K-fold cross validation is usually used to confirm the stable of model and prevent the model to be overfitted to training data. In K-fold cross validation, the whole dataset is split into equally K parts, and for each evaluation process different K-1 parts will be selected as training data and the rest part is the validation data. The error could be calculated as $E_k = \sum_{i \in k}(y_i - p(x_i))$ where $p()$ is the prediction of the model, k refers to the k-th part of evaluation data. Summarizing all errors, we could produce the cross-validate error as shown in equation 3.2.

$$CV = \frac{1}{K} \sum_{k=1}^{K} E_k \tag{3.2}$$

3.4.3 Hit-rate

The Hit-rate function is a special evaluate metric for Recommendation System. It calculates how many hits it predicted in TOP-N recommendation.

$$HR@n = \frac{num_{hits}}{n} \tag{3.3}$$

Equation 3.3 describes the hit-rate function for top-n recommendation. The number of hits refers to the predict item exists in real shopping list.

3.5 Random Forest with Imperfect Pairwise Data

As describe in Chapter 2, Online Shopping environment contains a lot of imperfect noisy data, in order to handle those kind of data we developed a Random Forest based algorithm and evaluate it with previous algorithms.

To address the problem, we firstly investigate the quality of information in Online Shopping context and conclude that we can not directly get well-organized clean data from Online Shopping environment. On the contrary, we could get a small set of well labeled data and a large set of noisy pairwise data. The pairwise data means all images organized in two pairs, we only know the similarity of them, and the noisy data means the label of that data could be wrong.

Based on the problem, we review previous pairwise and semi-supervised learning algorithms and find previous related work. However, all those works can not handle all problems(noisy,pairwise data,well-labeled data) at the same time. And based on previous research, we develop Pairwise Constraint Random Forest algorithm to solve the problem we addressed.

3.5.1 Data Collection and Experiment Design

In order to evaluate our methods with previous algorithms, we need to build attribute level dataset with both well-labelled data and noisy data. Meanwhile, to make our results more convincing, we prefer to use well-known and public data source rather than collect data by ourselves. In this case, we select three different public known datasets, the Clothing with Attribute dataset (Farhadi et al., 2009), DeepFashion Dataset (Z. Liu, Luo, Qiu, Wang, & Tang, 2016) and Animal with Attributes dataset (Lampert, Nickisch, & Harmeling, 2009). The details of datasets are described in Table 3.1. The categories in the table refer to the high-level class, which are typically the target of Machine Learning methods, for example, the category in clothing dataset contains suit, T-shirt, dress, and so on. The attributes are middle-level features such as colour, pattern and style, as described in Chapter 2, they are hard to classify than high-level categories. The number of attributes per image describes the coverage of the attribute label. For example, in DeepFashion dataset, there are totally 100 types of attributes, but not every image contains all attributes, on the contrary, the mean number of attributes per image is only 37.1.

Table 3.1: Features in three attribute level datasets

Datasets	#images	#categories	#attributes	#attributes per image
CwA dataset (Farhadi et al., 2009)	1846	5	26	26
DF dataset (Z. Liu et al., 2016)	289,222	50	100	37.1
AwA dataset (Lampert et al., 2009)	30475	50	85	81.7

Figure 3.2: Example from Clothing with attribute dataset

Figure 3.3: Example from DeepFashion Dataset

Figure 3.4: Example from Animal with Attribute datasets

The example of images for each dataset is shown in Figure 3.2, 3.3 and 3.4. As all data in these three datasets are well labeled, we randomly select part of them as well labeled one and then add noise label to the rest data. The advantages of choosing these datasets are in two folds: on the one hand, they are both well build data and has been used in many other studies which reduce the influence of bias information from untrusted data (Farhadi et al., 2009; Lampert et al., 2009; Z. Liu et al., 2016). On the other hand, by controlling noisy data manually, we can test our algorithm in different noisy rate which clearly experiment the robustness of algorithm. These two datasets are both public and free for educational usage.

In experiment development, we use C to implement SVM based algorithm, matlab to implement Random Forest based algorithm, python with TensorFlow to implement CNN based algorithm. All these implementations are based on open source softwares. By testing different types of algorithms in selected datasets, we use predict accuracy to describe how each algorithm works and get the conclusion of the Pairwise Constraints random forest works better than other existing solution.

3.6 Recommendation with Self Identity Theory

In this experiment, we try to enhance the Recommendation System with information generated by Computer Vision, then organize and analyze the system with Self Identity Theory. As a result, we introduce a unique Self Identity theory guided Recommendation system, and we also analyze the shopping behavior with Self Identity theory.

we reviewed the currently Recommendation System from both computer science and information system perspectives. From the review, we address the problems of the current research are the new Recommendation System developed by Computer Science researchers mostly focus on involve new features and get better accuracy, but they do not consider the feedbacks from customer and the evaluation do not include shopping history and customer behavior. And researches in information system area do not consider the new technologies and most research only consider recommendation as a factor of whole systems rather than concentrate on the Recommendation System. To solve the problem, we review Self Identity theory and introduce it can be used to guide and describe recommendation results. Based on the review, we build a new recommendation system with Self Identity theory, and use Self Identity theory to describe the shopping behavior of customers.

3.6.1 Data Collection and Experiment Design

The experiment is mostly focused on fashion items, so the data we collected are mostly fashion related. In order to analyze how Self Identity affected by different product's categories, we also collect a small set of data from digital items which compared with fashion items are easy to analyze. As described above, the whole framework contains three parts, attribute classifier, recommendation system and the Self Identity analyzer. To build the whole framework, we need three different types of dataset for training models and evaluate results. The datasets we used in the three stage are described in Table 3.2 .

The reasons we selected those datasets are firstly we need sufficient data to build the fashion item attribute classifier for image feature extraction. In this part, we use DeepFashion datasets (Z. Liu et al., 2016). Secondly, to build the recommendation system, we need customer shopping history data. Thus, we use Amazon digital public data as the base data and collect product details from Amazon.co.uk. Finally, as the most important part of the system, to analysing the recommendation system with Self Identity theory, we use click-through data from Amazon to analyse customer behaviours.

In experiment, we use python with TensorFlow to extract attribute level features, we use python to store and processing recommendation models. To

Table 3.2: Datasets for Recommendation and Self Identity theory

Stage	dataset	features of dataset
Attribute classifier	DF dataset (Z. Liu et al., 2016)	Attribute label, Fashion related
Recommendation System	Amazon dataset (R. He & McAuley, 2016)	Fashion related, item-customer relationship
Self Identity analyzer	Amazon dataset (R. He & McAuley, 2016)	Fashion related, time sequence data

analyzing the result we use python with Scikit-learn to show the distribution of different algorithms and different features.

3.7 Summary

By analyzing possible methodology in both Computer Science and Information Science, we select mathematical modeling and proof method and simulation method as our research methodology. We also introduce the protocol of our experiment and the associate evaluation metric for different problems. Finally, we introduce the dataset we used in our experiment and explain why we choose them.

Chapter 4

Improving Attribute Classification with Imperfect Pairwise Constraints

As described in Chapter 2, Semantic attributes extracted from images could help to improve Online Shopping Experience by automatic image classification and recommendation systems. However, learning of such attributes requires a large well-labeled dataset which is usually difficult and expensive to collect and sometimes requires human domain experts to annotate. Partially labeled data, on the contrary, are relatively easy to obtain from social media websites or be annotated by less experienced people. However, a partially labeled dataset usually contains a lot of noisy data which are challenging for previous methods. In this research, we propose a semi-supervised Random Forest algorithm that can handle a small well-labeled attribute dataset and large scale pairwise data at the same time for classifying grouped attributes. Results on two typical attribute datasets show that the proposed method outperforms a state-of-the-art attribute learner.

4.1 Introduction

Semantic attribute learning attracted a considerable growth of interest from computer vision and multimedia researchers in the last few years. By using those mid-level features, the accuracy of high-level category classifiers could be improved and the learned attributes can be adopted as semantic features for other multimedia applications. For example, using Computer Vision to recognize and describe fashion items at the semantic level is extremely helpful for analyzing personal fashion styles (Hu et al., 2015b). However, the training of a semantic model needs a large set of well-labeled data (Bossard et al., 2013; Di et al., 2013b)and as the fashion items increase every year, there is also a strong demand to update the current model with new data. The previous work mostly generated their datasets with the help of crowdsourcing (H. Chen, Gallagher, & Girod, 2012b; S. Liu et al., 2012) which is time consuming and expensive. So it is important to design attribute learning algorithms that do not require lots of well-labeled data.

One possible solution is transfer learning (Pan & Yang, 2010) or domain adaptation (Ben-David et al., 2010) which converting well trained models from a previous well-labeled dataset to the new dataset with limited labeled data. However, the limitation of both transfer learning and domain adaptation is they require the well-labeled and new dataset are related. This requirement is easy to fit when only considering category level information, but in this paper we mainly concentrate on attribute level categorization which are mostly dataset specified (e.g. dressing style in clothes dataset or sociality in animal dataset). Therefore, we considering another possible solution which is to make use of unlabeled and partially labeled data which can be easily obtained from the Internet or quickly annotated by humans. Utilizing both well-labeled and unlabeled/partially labeled data is a typical semi-supervised learning scenario. However, partially labeled data usually contains a lot of noisy information, especially for abstract attributes such as the style of clothing or the design of fashion items.

There are many existing approaches on semi-supervised learning. Some consider it as a clustering problem without taking advantage of the well-labeled data at all (de Amorim, 2012; Hu et al., 2008; Wagstaff et al., 2001; Zeng & Cheung, 2012; X. Zhu et al., 2015) and others use a well labeled dataset along with totally unlabeled data (X. Liu et al., 2013) but ignore the weak relationship in the unlabeled data. A most related solution to our work was presented by Nguyen & Caruana (Nguyen & Caruana, 2008) and Yan et al. (Yan et al., 2006) who used a well labeled dataset together with additional partially labeled data, but their methods do not work well with noisy data.

This research aims to provide a solution to solve the problem of learning attributes from partially labeled noisy data together with well-labeled data. The contribution of our work is two-fold. Firstly, our model is built on a small set of well-labeled training data with a large amount of noisy pairwise constraint data. The setting is more realistic than learning from well labeled data and can be easily collect in real online shopping environment. Secondly, the pairwise data is collected according to semantic groups (e.g. color or texture) and only two pairwise constraints (i.e. must-link, meaning a pair of samples must be in the same class, and cannot-link, meaning a pair of samples must come from different classes) are applied. This situation is shown in Figure 4.1, there are must-links in color, style, but cannot-link in neckline type. The purpose is to utilize such pairwise data and detect and reduce the noisy information to improve semantic attribute transfer learning.

Figure 4.1: Example for pairwise constraint data

4.2 Related work

The related works include three aspect researches, semantic attribute analysis, semi-supervised learning and transfer learning.

4.2.1 Semantic attribute analysis

As described in Chapter 2, attribute learning as a specific image classification problem has gained lots of interest and resulted in many successful applications Bossard et al. (2013); H. Chen et al. (2012b); Di et al. (2013b); S. Liu et al. (2012). Such approaches usually use low-level features to train lots of mid-level classifiers, and then the mid-level attribute information can be used to generate a more accurate high-level classifier or to analyze the semantic information of images. However, the training of mid-level classifiers relies on well-labeled data which are hard and expensive to collect. Furthermore, the selected attributes are mostly simple ones such as color or texture which do not contain much expert domain knowledge. Thus, the use of partially labeled data rather than well-label data makes the previous work more realistic.

4.2.2 Semi-supervised learning with pairwise constraints

We introduce supervised and semi-supervised learning in Chapter 2, and discussed how it can help with Online Shopping Environment. In this research, we experiment most learning methods described in Chapter 2 and particularly focussed on robust semi-supervised learning.

Semi-supervised learning is a class of machine learning techniques that makes use of unlabeled or partially labeled data for training. In this paper, we specify the partially labeled dataset as pairwise constraint data and analyze the existing work which can handle pairwise constraint relations.

To utilize pairwise constraints, some researchers consider it as a clustering problem and develop constrained clustering algorithms based on K-means such as COP K-means Wagstaff et al. (2001) and CMWK-Means de Amorim (2012). With a similar idea, a more efficient Semi-supervised Maximum Margin Clustering method was also introduced to handle the pairwise constraints Hu et al. (2008); Zeng and Cheung (2012). The drawback of these algorithms is that they make a strong assumption that the pairwise constraint data are accurate. However, in practice, pairwise information is usually collected from social media sites which makes it contain a lot of noisy data. To avoid the influence of noisy data, the Constraint Propagation Random Forest which attempts to prevent the noise impact by an ensemble is formulated X. Zhu et al. (2015). Compared to the previous algorithms, it performs better when dealing with noisy data. However, the limitation of these algorithms is that they only use the pairwise constraint data to do clustering rather than classification. To build attribute classifiers, a small set of well-labeled data should be utilized along with a large amount of pairwise data.

Taking labeled data into consideration, Liu et al X. Liu et al. (2013) proposed a Random Forest based approach which trains trees with both labeled and unlabeled data, and as an ensemble method it is also robust to noisy data. However, this algorithm does not take the relation among unlabeled data into consideration, but only uses them as background knowledge to find better splits. To use labeled data and pairwise information at the same time, Convex Pairwise Kernel Logistic Regression which builds a loss function with pairwise constraint information was introduced by Yan *et al*let@tokeneonedot. Yan et al. (2006). Similar with this work, a Margin-based approach was also proposed Nguyen and Caruana (2008). In this method, a regular multi-class SVM is extended with a pairwise optimization objective function. These two algorithms use both well-labeled and partially labeled data, but according to our experiments, they are not robust to noisy pairwise data and do not consider the different qualities between well-labeled data and pairwise data. In addition, based on Spectral Kernel Learning, Shang*et al*let@tokeneonedot. Shang, Jiao, and Liu (2012) proposed a semi-supervised classification algorithm with enhanced spectral kernel under the squared loss (ESKS) which also takes labeled data and pairwise labeled data into consideration. However, the target dataset of ESKS is with large labeled data and small number of unlabeled or partially labeled data. That is different with the situation we considered in our work which the dataset contains small number of labeled data and large number of partially labeled data.

Another group of studies taking unlabeled data into their model and proposed some more weak supervision situations (Fogel, Averbuch-Elor, Cohen-Or, & Goldberger, 2019; Ying et al., 2018). However, because of the difficulties of their proposed data settings, they only focused on clustering problem rather than the attribute level classification. Furthermore, there are still some works trying to do difficult classification problem based on only unlabeled and partially labelled data (Shi, Otto, & Jain, 2018; Tu, Lin, Wang, & Kim, 2018). Nevertheless, those works only focused on a specific research area such as Facial recognition and digital signal processing, and they put much prior knowledge on their algorithm design which makes it difficult to be extended to other types of datasets. As we discussed above, the previous algorithm can not solve the robust pairwise constraint attribute learning problem in the particular data setting proposed in this thesis.

4.2.3 Transfer Learning

Deep learning achieves considerable success in most Machine Learning problems, especially in the Computer Vision field. However, Deep learning model usually needs many data to train, which is hard and expensive in most online shopping environment (Schmidhuber, 2015). Transfer Learning which transfers a well-trained model from one domain to another solves the insufficient data problem, but it requires the different domains are related (Pan & Yang, 2010). For recognizing new attributes in new datasets, the transfer learning still needs enough data to cover new features (Weiss, Khoshgoftaar, & Wang, 2016). The recent research for transfer learning is mainly focused on domain transfer problem and try to introduce new techniques from other research fields to accelerate the training for new model (C. Tan et al., 2018). For example, Cui Ying et tl (Cui, Song, Sun, Howard, & Belongie, 2018) proposed a domain transfer training strategy for fine-grained categorization. Moreover, to make the domain transfer more accuracy, a partial domain transfer algorithm which transfers from a big domain to a small sub-domain were also introduced (Cao, Long, Wang, & Jordan, 2018). Another research tried to solve the lack of data problem for transferring, they use self-supervised pretraining to make use of unlabeled data (Noroozi, Vinjimoor, Favaro, & Pirsiavash, 2018). However, considering the problem proposed in our thesis, the major problems are not domain transfer or data enrichment but to use the information from pairwise constraint links, and the previous research is helpful but not that relative to our tasks.

4.3 Robust Random Forest with Pairwise Constraints

As mentioned in the introduction, pairwise constraints exist among the related attributes, so in our approach, we split different attributes into various groups. As a result, the pairwise relation only exists in the same group. To address the problem clearly, the following Figure 4.2 describes the overall flowchart of the proposed Robust Random Forest with Pairwise Constraints. Compared with regular Random Forest, the new algorithm contains a pairwise node splitting strategy discussed in Section 4.3.2 to handle pairwise data, and a tree evaluation method described in Section 4.3.3.

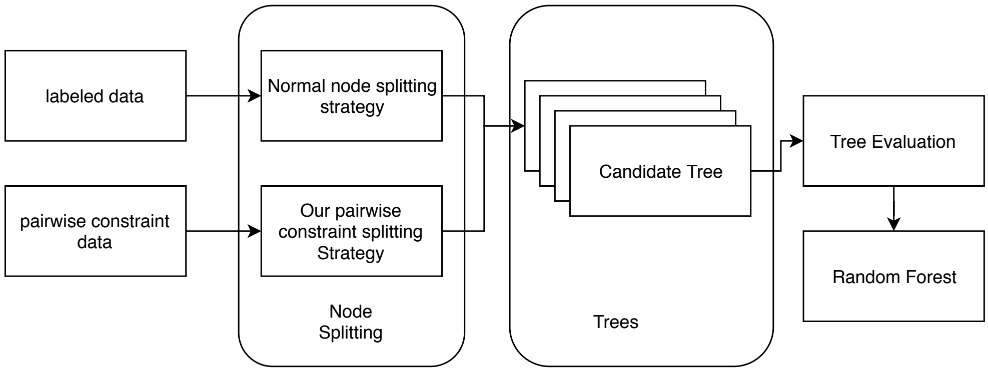


Figure 4.2: Flowchart of Robust Random Forest With Pairwise Constraints

A fully supervised dataset usually includes a set of labeled training samples $L = \{x_i \in X\}_{i=1...l}$ and their labels $\{y_i \in Y\}_{i=1...l}$ where $X \subset \Re$ is the feature space, $Y = \{1...k\}$ is the label set and l is the number of samples. In this paper, in addition to the labeled data, there are also partially labeled data which include Must-Link $M = \{(x_i^\alpha, x_i^\beta) \| y_i^\alpha = y_i^\beta\}_{i=1...m}$ and Cannot-Link $C = \{(x_i^\alpha, x_i^\beta) | y_i^\alpha \neq y_i^\beta\}_{i=1...n}$. In our learning framework, different with the previous one proposed by Nguyen & Caruana Nguyen and Caruana (2008), we consider a more realistic situation where the partially labeled data contain a lot of noisy information. And the labeled dataset, on the other hand, is assumed to be a well-constructed one with fewer noisy data.

4.3.1 Random Forest with Supervised Learning

Random Forest is a widely accepted ensemble method to handle noisy data Breiman (2001). It consists of a list of decision trees$\{t_1, t_2, ..., t_N\}$ which are independently trained with a random subset of the whole data and the final results are generated by getting votes from all these trees. Growing each decision tree involves data selection, node splitting and stopping criterion detection. According to a predefined sub-sample rate r, the data selection phase is to randomly select a subset from the whole dataset Ho (1998). Feeding the selected data into decision trees, each node splits the data into two parts with a splitting strategy. Considering both accuracy and efficiency, the Linear Combination Splits Geurts, Ernst, and Wehenkel (2006) is the most used one. The function of it is defined as follows:

$$h(W, \theta) = \begin{cases} 0, W \cdot x < \theta \\ 1, otherwise \end{cases} \tag{4.1}$$

where W is the parameter with one or more non-zero elements to select features, and θ is the split threshold. According to the output of $h(W, \theta)$, all arriving samples are split into either the left or right child node . To find the optimal split parameter, the criterion ΔG is introduced and it can be formulated as

$$\Delta G(R) = G(R) - \frac{|R_l|}{|R|}G(R_l) - \frac{|R_r|}{|R|}G(R_r) \tag{4.2}$$

where R refers to the current node, R_l, R_r represent the attempted split to left and right child nodes. The function G can be computed by many methods such as information gain and Gini impurity Olshen, Stone, et al. (1984). In this paper, we choose Gini impurity $\sum_{i \neq j} p_i p_j$ due to its efficiency. The p in Gini impurity reflects the proportion of samples belonging to the same category, and it can be calculated as follows:

$$p_l = \frac{1}{|R|}\sum_{i=1}^{|R|}[y_i = l] \tag{4.3}$$

where $|R|$ is the number of samples in current node R and $[y_i = l]$ refers to indicator function,it equals to 1 when $y_i = l$ and otherwise equals to 0.

The target of information gain is to maximize ΔG by selecting different W and θ. To make it more efficient, a maximum attempting number m_{try} is usually defined as a stopping criterion.

After training, each tree in Random Forest gives a probability estimation of class l, $p_t(l|x)$ for a given test case x, and then the total probability of random forest is calculated by averaging

$$p(l|x) = \frac{1}{N}\sum_{i=1}^{N} p_i(l|x) \tag{4.4}$$

where N is the number of trees and $p_i(l|x)$ is obtained by calculating the ratio of class l getting votes from the leaves in the ith tree. The final result of Random Forest is defined as

$$\hat{l} = \underset{l \in Y}{argmax}\; p(l|x) \tag{4.5}$$

4.3.2 Node Splitting with Pairwise Constraints

The conventional Random Forest described above only takes labeled data as input and the limited size of supervised data in our problem would lead to an obvious performance drop X. Liu et al. (2013). To avoid this problem, we extend the current splitting strategy to take pairwise constraint data into training. According to equation 4.2, the target of split is to maximum the Gini impurity at each node which requires to obtain the proportion of well labeled data belonging to the same category. However, for partially labeled data, samples come in pairs, so when the Must-link or Cannot-link relation is broken by splitting, the gini index calculated by $\sum_{i \neq j} p_i p_j$ does not work. In this situation, we introduce a new method for both Must-link M pairs and Cannot-link C pairs. In pairwise constraint situation, instead of directly calculating the number of samples falling into the correct side of each node, we count the ratio of broken links in total must-link and cannot-link. So the following equation will illustrate how we calculate that ratio.

Equation 4.6 obtains the number of samples in Must-link set M falling into the same node, and Equation 4.7 calculates the total number of samples from Must-link set M in the node R.

$$N^M(R) = 2 * |\{(x^\alpha, x^\beta)|x^\alpha \in R \wedge x^\beta \in R \wedge (x^\alpha, x^\beta) \in M\}| \tag{4.6}$$

$$\begin{aligned} N^M_{total}(R) = &|\{x^\alpha | x^\alpha \in R \wedge (x^\alpha, x^\beta) \in M\}| \\ &+ |\{x^\beta | x^\beta \in R \wedge (x^\alpha, x^\beta) \in M\}| \end{aligned} \tag{4.7}$$

where R denotes the current tree node. α and β refers to the pair data in a pairwise constraint. The reason we twice the result in the final Equation 4.6 is because there are two samples in a pair. Similar to Must-link, we obtain the same data from Cannot-link set C with Equation 4.8 and 4.9

$$N^C(R) = 2 * |\{(x^\alpha, x^\beta)|x^\alpha \in R \wedge x^\beta \in R \wedge (x^\alpha, x^\beta) \in C\}| \tag{4.8}$$

$$\begin{aligned} N^C_{total}(R) = &|\{x^\alpha | x^\alpha \in R \wedge (x^\alpha, x^\beta) \in C\}| \\ &+ |\{x^\beta | x^\beta \in R \wedge (x^\alpha, x^\beta) \in C\}| \end{aligned} \tag{4.9}$$

The Equation 4.6 , 4.7 and 4.8, 4.9 calculate the number of pairwise constraint samples falling into same side or different side and the total number of must-link and cannot-link. That information will be used further to calculate the radio of success splitting for pairwise constraints.

Based on N and N_{total} defined above, we propose two estimation functions as follows:

$$E^M = -\log \frac{N^M(R)}{N^M_{total}(R)} \tag{4.10}$$

$$E^C = -\log \frac{N^C_{total}(R) - N^C(R)}{N^C_{total}(R)} \tag{4.11}$$

Equations 4.10 and 4.11 are the estimation functions for Must-link and Cannot-link sets respectively. These two equations are constructed by the ratio between successful splitting number and total samples. In addition to making it more robust for noisy data, we apply $\log$ function to it, which makes derivative of it smaller when the successful number is close to the total number. Actually, there are other ways to define the estimation function, but in our experiment, this is the most simple and efficient solution.

To use the estimation proposed above for node splitting, this project follow the idea from the origin Random Forest as discussed in Equation 4.2, we can propose a similar split criterion for pairwise constraint data which evaluating the tree before and after splitting.

$$\begin{aligned} \Delta E(R) = E^M(R) - \frac{|R^M_l|}{|R^M|}E^M(R^M_l) - \frac{|R^M_r|}{|R^M|}E^M(R^M_r) \\ + E^C(R) - \frac{|R^C_l|}{|R^C|}E^C(R^C_l) - \frac{|R^C_r|}{|R^C|}E^C(R^C_r) \end{aligned} \tag{4.12}$$

Then the new target becomes to find a split to maximize Equation 4.12 which contains both must-link estimation and cannot-link estimation.

In addition, our algorithm considers a more complex condition that the pairwise data include a lot of noise information, when the procedure of tree construction closes to the leaf nodes, the total number of pairwise data $N^M_{total}(R)$ and $N^C_{total}(R)$ could be smaller and Equation4.10 and 4.11 would too sensitive to noise data. In order to avoid this problem, we introduce a combined split strategy as follows:

$$\Delta C(R) = \begin{cases} \Delta G(R) + \alpha \Delta E(R), |L| < |M| + |N| \\ \Delta G(R), otherwise \end{cases} \tag{4.13}$$

where $|L|, |M|, |N|$ refers to the number of well-labeled, must-link and cannot-link data respectively, and α is the learning rate for pairwise data. In practice, we usually choose a small number for α which makes $\Delta E(R)$ only have a limited influence at tree construction.

In this paper, we assume a situation that the size of partially labeled data is much larger than that of the well-labeled data, but the accuracy is on the contrary. So to calculate Equation 4.13 more efficiently, the samples with broken Must-link or satisfied Cannot-link will be removed from child nodes as shown in Equations 4.14 and 4.15. With the help of this data filtering strategy, the number of total pairwise data will be reduced in child node and make the algorithm more efficiency.

$$M^{new}(R) = \{(x^\alpha, x^\beta) | x^\alpha \in R \wedge x^\beta \in R \wedge (x^\alpha, x^\beta) \in M\} \tag{4.14}$$

$$C^{new}(R) = \{(x^\alpha, x^\beta) | x^\alpha \in R \wedge x^\beta \in R \wedge (x^\alpha, x^\beta) \in C\} \tag{4.15}$$

The tree constructing procedure is shown in Algorithm 1. It illustrate how our pairwise constraint node splitting can be merged into normal random forest. When split strategy set to $\delta C = \Delta G(R)$ it works as normal random forest, otherwise it works based on pairwise constraint information.

4.3.3 Evaluation of Trees with OOB Error

In our work, we consider the partially labeled data contain a lot of noisy information, and the Random Forest, as a ensemble method, reduces the influence of noisy data by aggregating results from the individual trees. However, the splitting strategy described above considers all the data from both labeled and partially labeled datasets, which means the noisy data still have contributions to the result of each tree. Therefore, when the noise rate of partially labeled dataset is large, most of the trees in Random Forest will be affected by the wrong information and the final result from Random Forest could be less accurate. So there need to be an evaluation method to select training samples.

The training of each tree in the Random Forest is independent and only uses a subset of training data. The samples which are not used for training are defined as the Out Of Bag (OOB) data Breiman (2001) and the error calculated from it is known as an unbiased estimation of the accurate of trees. Therefore, to reduce the influence of noisy data in a partially labeled dataset, each tree is trained on only the labeled data or both labeled and partially labeled data. If the OOB error increases with the tree trained on both datasets,

Algorithm 1: Construct Random Forest Tree

Input : Well-labeled training data $X_l \in L$ in current node R, pairwise data $(X^\alpha, X^\beta) \in \{M \cap C\}$ in current node R,learning rate for pairwise data α ,maximum splitting attempting numberm_{try}

Output: The best cut hyperplane W and θ, The associated child node R_l and R_r

Remove unnecessary pairwise training data according to Eq4.14 and 4.15;
if $|x_l| < (X^\alpha + X^\beta)/2$ **then**
  Set split strategy $\delta C = \Delta G(R) + \alpha \Delta E(R)$;
end
else
  Set split strategy $\delta C = \Delta G(R)$;
end
Initialise W, θ;
$n_{try} \leftarrow 0$;
$\delta C_{best} \leftarrow 0$;
repeat
  $n_{try} = n_{try} + 1$;
  Randomly select W_0 and θ_0;
  calculate δC according to selected split strategy;
  if $\delta C_{best} < \delta C$ **then**
    $\delta C_{best} \leftarrow \delta C$;
    $W \leftarrow W_0, \theta \leftarrow \theta_0$;
  end
until $n_{try} > m_{try}$;
Generate R_l, R_r according to $W and \theta$;
return $W, \theta and R_l, R_r$;

these two trees will be both deleted and new sub-samples of the partially labeled dataset will be randomly generated for training new trees. With the help of this method, a subset which contains too much noisy data will be ignored and a new one will be generated. And this procedure can be used to detect and ignore noisy data. The tree selecting procedure is shown in Algorithm 2. There are two hyperparameters defined for this algorithm, the maximun training times and bootstrap aggregating ratio. In our experiment, the best values for each hyperparameters are determined by KFold evaluation and in most case, we set aggregation ratio to 0.6 and maximum training times to 50.

Algorithm 2: Select Random Forest Tree

Input : Well labeled training data $X_l \in L$, pairwise data $(X^{\alpha}, X^{\beta}) \in \{M \cap C\}$, maximum training times n_{max}

Output: Random Forest Tree T

$X_l^i \leftarrow$ generate a new subset from X_l using bootstrap aggregation;
Training tree T_l with X_l^i;
$oobe_l \leftarrow$ Compute the OOB error of T_l;
$n_{try} \leftarrow 0$;
$best_{oobe} \leftarrow oobe_l$,$best_T \leftarrow T_l$;
repeat
    $n_{try} = n_{try} + 1$;
    $(X^{\alpha}, X^{\beta}) \leftarrow$ generate a new subset from M and C using bootstrap aggregation;
    Training tree T with both X_l^i and (X^{α}, X^{β});
    $oobe \leftarrow$ Compute the OOB error of T;
    if $best_{oobe} > oobe$ **then**
        $best_{oobe} \leftarrow oobe$,$best_T \leftarrow T$;
    end
until $n_{try} > n_{max}$;
return $best_T$;

4.4 Transfer Learning with imperfect pairwise data

Deep learning usually outperforms shallow learning methods in most Computer Vision problems, but it requires a lot of data to train the model. Applying transfer learning to deep model may reduce the requirement for the number of data, but for complex and professional target such as attribute learning, a large number of data is still required.

In our data setting, the transferred model needs to learn new attributes from the imperfect pairwise data and well labeled data at the same time. As a result, the normal Sigmoid or Cross-entropy loss does not fit for this situation. To handle this problem, we introduce triple loss into our model as shown in equation 4.16.

$$L_{x_a,x_p,x_n} = \sum_{j=1}^{|X|} max(\alpha + \|x_i^a - x_i^p\|_2^2 - \|x_i^a - x_i^n\|_2^2) \tag{4.16}$$

where the input is a triplet (x^a, x^p, x^n), where x^a refers to anchor data which come from the well-labelled dataset, and x^p, x^n refers to the positive and negative samples generated by pairwise data. x^p, x^n is different from the pairwise pair x^α, x^β discussed above, which the new samples have been already identified whether it is similar or dissimilar to the well-labelled data.To reduce the noisy data in pari-wise set, we train a PCRF firstly and use the Evaluation of Trees described in Section 4.3.3 to ignore the noisy pairs and use the rest data for training.

The deep learning model will be firstly trained on large dataset(e.g. ImageNet (Deng et al., 2009)) and then keep the top level layers fixed and transfer learn the well labeled data for bottom layers. Finally the deep model will be fine-turned on cleaned pair-wise data with triplet loss.

4.5 Experiments and Analysis

To test the performance of Pairwise Constraint Random Forest, we build two different attribute level datasets based on the Clothing with Attributes dataset Farhadi et al. (2009) and Animal with Attributes dataset Lampert et al. (2009). Considering the efficiency, the algorithm proposed in this paper is implemented based on the code provided online[1].

4.5.1 Dataset settings

Clothing with Attributes dataset

The Clothing with Attribute dataset includes 11 different attribute groups and totally 36 attributes. According to the learning framework described above, a small labeled dataset and a large pairwise dataset should be both generated. To make the pairwise dataset more realistic, we

[1] https://github.com/karpathy/Random-Forest-Matlab

choose 5 major colors as the color group, 6 different patterns as the pattern group, 3 different lengths of sleeve and 3 types of neckline shape as another two groups. Then we only keep 300 well labeled data and randomly generate pairwise datasets with 800 and 1500 images. To simulate the noisy situation, the four attribute groups are manually added noisy data with 0%, 5%, 10%, 15% and 20% respectively.

Because the Clothing with Attribute dataset only provides the original images, we apply the pose estimation method by Yang et al. Yang and Ramanan (2013) to locate the clothing area. And with the help of VLFeat Vedaldi and Fulkerson (2008) and scikit-learn Pedregosa et al. (2011), low level features including SIFT features, Color Histogram features, Local Self-Similarity features and PHOG features are extracted. The color histogram features are extracted from 21 cells of a 3-level spatial pyramids of images and for each cell 128 dimensional histogram are extracted and concatenated into 2688 feature vectors. For PHOG features, same spatial pyramids are applied to images but 12 dimension feature vectors are extracted for each cell and result in 252 feature vectors. The other features are extracted by 2000-bin bag of words histograms. Then we use PCA to reduce the high dimension features and keep 99% information. We do not apply PCA on PHOG features and the other results is 1359 for color histogram, 1875 and 1391 for SIFT and LSS separately.

Attribute Groups	Colors	Patterns	Length of sleeve	Neckline shape
Number of Attributes	5	6	3	3

Table 4.1: Chance of different data to be selected for training

Animal with Attributes dataset

For the Animals with Attributes dataset, there are 85 different attributes in total. In our experiment, we randomly select 5 attributes as the color group, 4 attributes as the texture group, 5 attributes as the living place group and finally 5 attributes as the sociality group. Because this dataset is much bigger and much nosier than the previous one, we select 400 samples as the well-labeled dataset and 1000 and 3000 samples as pairwise data. The pairwise dataset is also manually added noisy information at 0%, 5%, 10%, 15% ,20%. As the low level features are provided by authors Lampert et al. (2009), we directly use these features and the dimension of each features are same with our previous dataset. After applying PCA to them and maintain 99% information. The number of dimension of each feature vectors are 1538, 1797, 1546 for Color histogram, SIFT and LSS separately.

4.5.2 Reduce the influence of noisy data

Random Forest uses a subset of training data to build each tree. The sub-sampling rate is the number of samples in a subset divided by that in the whole dataset Breiman (2001); X. Liu et al. (2013). In traditional Random Forest, the sub-sampling rate is manually defined before training and the chance of samples to be selected into a subset is equal during training. After the sub-sampling,the data selected by each trees will all contribute to the evaluation for node splitting.

However, in Pairwise Constraints Random Forest, according to Algorithm 1 the pairwised data used in training could be ignored and the influence of them is also limited. And as we described in Section 3.3, to reduce the influence of noisy data, we evaluate each tree with OOB error and keep generating new subsets from the partially labeled dataset that means the randomly generated subsets which contain too many noisy data will be deleted. As a result, even if we fix the sub-sampling rate before training, the chance to choose noisy data into a subset is less than the chance to choose normal data. In the experiment, because the dataset is not balanced, we choose different

Noise rate	0%	5%	10%	15%	20%
Well-labeled data	90%	90%	90%	90%	90%
Partially labeled data	60%	60.60%	61.87%	63.57%	66.63%
Noisy data	60%	48.51%	43.20%	39.37%	33.48%

Table 4.2: Chance of different data to be selected for training

sub-sampling rates for labeled and unlabeled data as 0.9 and 0.6 respectively. Analyzing the procedure of Pairwise Constraint Random Forest working on both dataset, table 1 shows the average chance of different samples to be selected in all trees during training. The results indicate that the effect of noisy data is reduced by the methods we proposed in this paper.

4.5.3 Compared with other pairwise algorithms

There are limited previous works on classification with a small well labeled dataset and a large pairwise constraint dataset. As shown in Nguyen and Caruana (2008)'s research, the proposed PCSVM reached the state-of-the-art performance and PKLR proposed by Yan et al. (2006) works better when the data is limited. So we compare our pairwise constraint Random Forest with PCSVM and PKLR in the same dataset proposed above. The implementation of PCSVM and PKLR algorithm is provided by authors. We only make minor changes to make it work for our datasets. To illustrate the effect of partially labeled data, we test a normal Random Forest training with only well labeled dataset.

To get the best performance of each algorithm, we apply 5-fold cross validation to select best parameters. For the pairwise constraint Random Forest we set the number of trees to 150 and the learning rate for partially labeled data is set to 0.01.

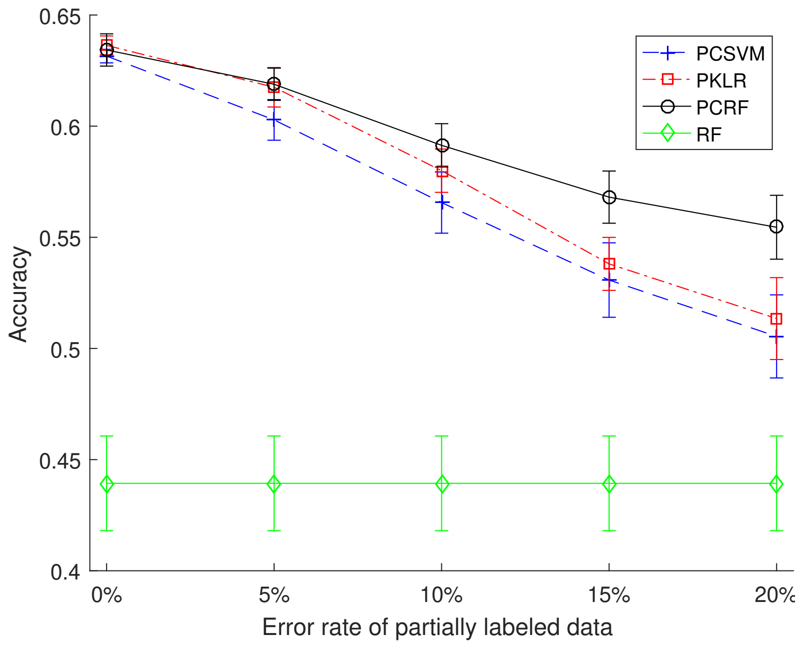


Figure 4.3: Results of PCRF on large Clothing with Attribute dataset

The dataset used in our experiments contains only limited well-labelled data and a lot of noisy information, so the stability of the algorithm is important. To get the variance of different methods, we apply 5-fold validation on testing and choose the average accuracy along with variance as the results. Figure 4.3, 4.4, 4.5, 4.6 shows the average performance of all attribute groups with Pairwise Constraint RF against PCSVM, PKLR and the normal Random

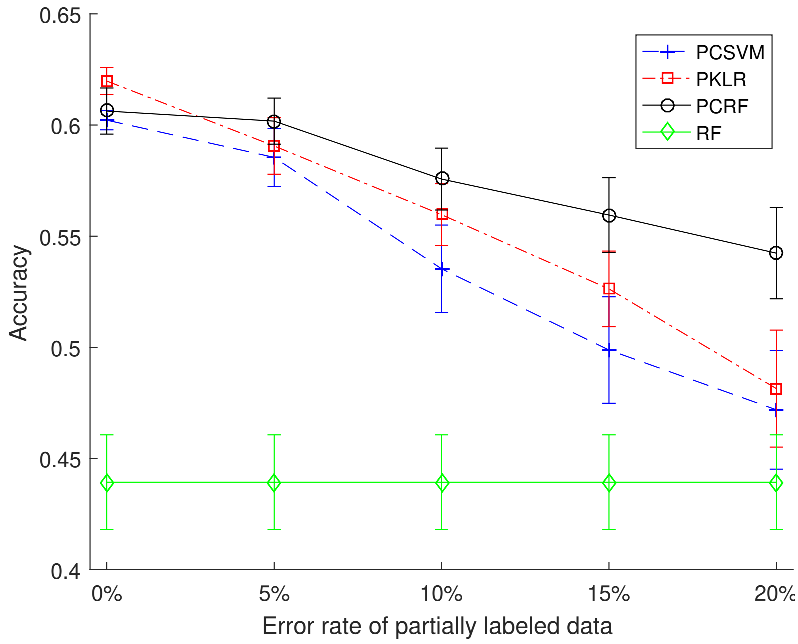


Figure 4.4: Results of PCRF on small Clothing with Attribute dataset

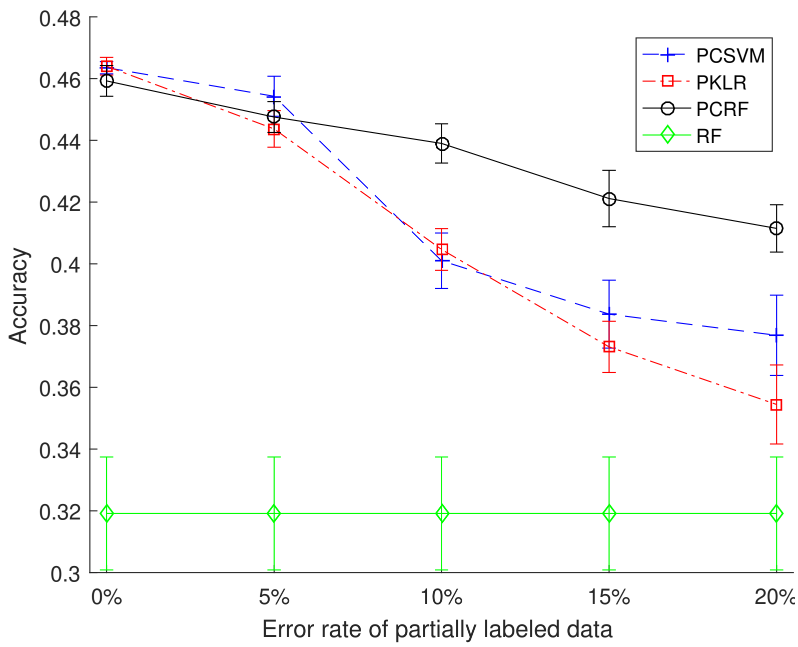


Figure 4.5: Results of PCRF on large Animal with Attribute dataset

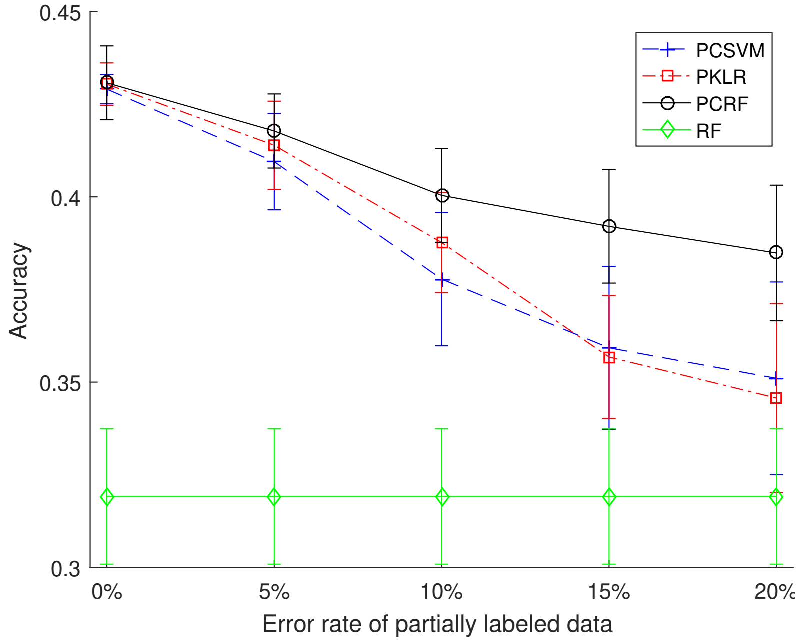


Figure 4.6: Results of PCRF on small Animal with Attribute dataset

Forest. It is clear that when the noise rate is lower, all algorithms get a similar accuracy, which means all algorithms work well with accurate pairwise constraint data. However, when the noise increases in the pairwise data, Pairwise Constraint Random Forest is more robust and performs better than PCSVM and PKLR.

As demonstrated in the results we get from a small set and large set partially data in Figure 4.3,4.4 and 4.5, 4.6, Pairwise Constraint Random Forest can handle both small and large set of partial data, and the variances of results on both datasets are smaller than others. Meanwhile, the increasing number of partially labelled training data usually results in a more robust model. This implies that the sufficient pairwise labelled data is important for getting a robust model, but the accuracy is limited due to the ability of random forest.

Our model outperforms others when the dataset contains a lot of noise data mainly because the particular tree selection strategy can reduce the influence of noise data and the random forest based algorithm itself is robust to noise information. However, the accuracy still suffers from the increasing noise rate, which means our algorithm still cannot avoid the influence of noise data completely.

Figure 4.7 and 4.8 shows the average performance of all attribute groups with PCSVM against Pairwise Constraint RF and the normal RF with only labeled data on both datasets. When the noise rate is smaller, PCSVM and Pairwise Constraint RF get the similar accuracy which means both algorithms

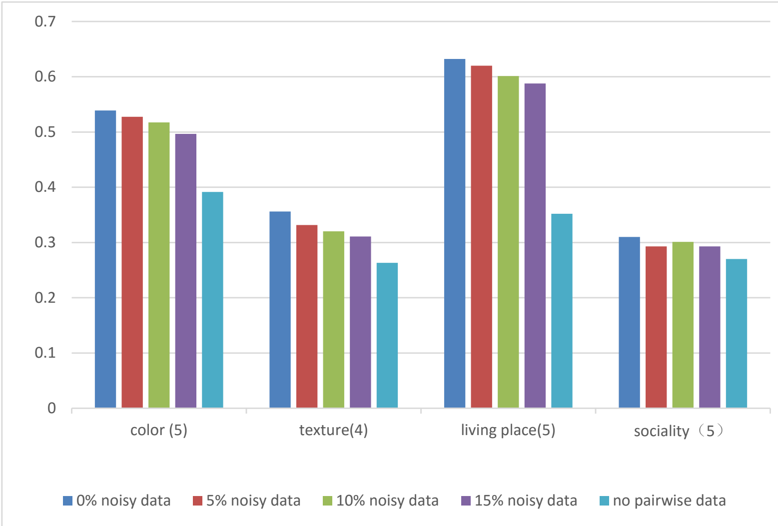


Figure 4.7: Average accuracies on Animal with Attribute dataset

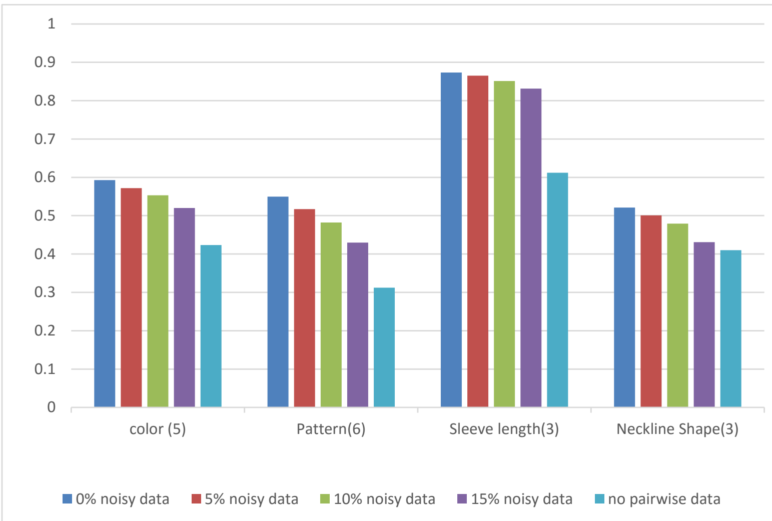


Figure 4.8: Average accuracies on Clothing with Attribute dataset

work well with accurate pairwise constraint data. However, when the noise increases in the pairwise data, Pairwise Constraint Random Forest is more robust and performs better than PCSVM.

The result indicate PCRF can successful reduce the influence of noisy data, but the accuracy of PCRF is still limited due to the quality of image feature and ability of random forest. To get a better solution, we introduce the deep learning based transfer learning model with PCRF algorithm to reduce noisy data.

4.5.4 Compared with Transfer Learning

Transfer learning Pan and Yang (2010) as an important field in Deep Learning shows great advantage in Computer Vision fields. However, transfer learning usually bases on a pre-trained model which is similar to the target dataset. And to fine-turning the pre-trained model, it still need some well-labeled data Weiss et al. (2016). In our experiment, to test how transfer learning works on our dataset, we selected ResNet50 model pre-trained on ImageNet and fine turning on the small set well-labeled data. Transfer learning performs similar in low level features such as Color and Pattern, but it does not works well in complex attribute features. The reason is in two fold, on the one hand, the pre-trained dataset is not similar with the target dataset. On the other hand, the well-labeled data is quite limited, the deep learning model needs more data to fit the new dataset. Since PCRF can efficient reduce the noisy of pair-wise training data and classify them into binary classification, we proposed a combination learning strategy including both transfer learning and PCRF. The noisy pair-wise data will firstly feed into PCRF where the noisy data will be removed, then we choose those classified data with high confidence as labeled data to fine turn the transfer learning model. The results illustrate in Table 4.3 and Table 4.4 proves the accuracy of our combined model.

The RF in Table 4.3 and Table 4.4 refers to normal random forest algorithm with only labeled data, similarly, TF-Labeled is the transfer learning model with only labeled data. Triplet loss is a special loss function used to handle pairwise data, TF-Triplet is the transfer learning model with both labeled and noisy pairwise data. TF-PCRF is the model we presented above which combine the advantage of transfer learning and PCRF, the results also prove the TF-PCRF is the state-of-the-art solution for such datasets.

Table 4.3: PCRF with Transfer Learning on Animal with Attribute dataset

Model	color	texture	living	place	All
RF	0.3915	0.2631	0.352	0.27	0.31915
PCRF	0.4792	0.3022	0.5784	0.2861	0.411475
TF-Labeled	0.5175	0.3102	0.3312	0.291	0.362475
TF-Triplet	0.5013	0.319	0.4132	0.2532	0.371675
TF-PCRF	0.5318	0.3421	0.5821	0.3268	0.4457

Table 4.4: PCRF with Transfer Learning on Clothing with Attribute dataset

model	color	Pattern	Sleeve length	Neckline Shape	All
RF	0.4233	0.312	0.612	0.41	0.439325
PCRF	0.5177	0.4264	0.827	0.447	0.554525
TF-Labeled	0.5861	0.3531	0.7231	0.4132	0.518875
TF-Triplet	0.5485	0.4172	0.837	0.4513	0.5635
TF-PCRF-Triplet	0.6029	0.5265	0.8572	0.5012	0.62195

In our experiments, the deep learning based TF methods outperform random forest based methods is basically because the deep learning can handle unstructured visual images better, the automatic generated deep features can represent images better than handcrafted features such as SIFT, colour histogram. Compared with TF-Labeled and TF-Triplet, it is clear that the use of pairwise constraints with the help of triplet loss can increase the training data and achieve better accuracy. And considering the noise data in pairwise constraint dataset, the joint using of TF PCRF with triplet can successfully reduce the influence of noise training information.

Due to the quality and quantity of our experiment datasets, the proposed TF-PCRF-Triplet model achieve best accuracies in all experiment tests, but the result is still far away from the accuracy of deep learning models on other famous datasets(e.g. ImageNet). Considering the difficulty of the attribute we want to predict and the quality of data we used, our work still produces an acceptable solution.

4.6 Conclusion

In this paper, we addressed the problem of the lack of well-labeled data for training attribute classifiers, grouped related attributes and proposed a Pairwise Constraint Random Forest which handles well labeled data and imperfect pairwise data at the same time. In the experiments, we extended two well-known attribute datasets with different noise rates and test the performance of Pairwise Constraint Random Forest on them. The results showed that the Pairwise Constraint Random Forest proposed in this paper works better than the previous methods on a large set of pairwise noisy data. Based on the result, we combine transfer learning and PCRF to get a more accuracy solution.

Chapter 5

Recommendation with Self Identity Theory

5.1 Introduction

With the explosive growth of data available on online shopping environment, customers are hard to find out their best choice from countless products. Recommendation systems(RS), as an information filtering tool, can deal with such problem by providing users with new contents and products they probably are interested in (Jannach, Zanker, Felfernig, & Friedrich, 2010). RS is a part of online shopping environment to work out customer's over-choice, and the research of it has been done in both Computer Science and Information System areas (Ricci et al., 2015). The aim of those researches are both to improve Online Shopping Experience, but Computer Science researches are mostly focused on algorithms which makes RS more accuracy and researches in Information System fields are more concentrate on customers and the whole systems that explaining how and why RS helps in Online Shopping Experience (Childers et al., 2002; Park & Kim, 2003) Therefore, how IS researches and theories can help to analyze and improve CS Recommendation Systems and how CS algorithms can improve the understanding of Online Shopping Experiences could be an interesting problem. In this experiment, we try to build an high explainable RS and using Self Identity theories from IS to analyze the information from RS. Based on the result, we finally introduce a better RS design with the support from IS theory.

In previous IS research, the RS is mostly considered as a black box in the

Online Shopping Environment to produce predictions and help customers to improve Online Shopping Experience (Hashim et al., 2009; Hernández et al., 2011; G.-G. Lee & Lin, 2005). There are rarely RS centered research and the reason is in IS we need to collect lots of theoretical related factors to support some theories and conclude results, but the current RS itself can not provides enough customer and product related data Ricci et al. (2015). In our research, we are trying to solving this problem by making RS more explainable and taking more complex data into RS.

In CS, the research of RS are mostly concentrated on taking use of new information and bring better Machine Learning algorithm into RS (Adomavicius & Tuzhilin, 2005). Generally, the recommendation results are generated based on personal preferences, product features, user-item relationship and time sequence information (Adomavicius & Tuzhilin, 2005; Ricci et al., 2015). The algorithms of RS can mainly categorized into collaborative filtering, content-based recommendation and hybrid recommendation (Adomavicius & Tuzhilin, 2005). Those traditional approaches can produce reasonable predicts but the results are mainly based on simple user-item relationships and features(text and label features), there are a lot of complex information(e.g. comments, images) in Online Shopping environment are not been used. When trying to analyzing the prediction in IS view, we found there are not enough customer and product factors to support any theories. The RS can only considered as a whole system or a black box in Online Shopping Environment. Recently, deep learning are widely used in many application domains such as Computer Vision and Nature Language Processing because of its capability in solving complex data types and providing state-of-the-art solutions (Lecun et al., 2015). Therefore, deep learning brings more opportunities to improve the performance of RS and there are already lots of deep learning based RS has been developed to take complex data into current recommendation architectures (Ricci et al., 2015). Those new deep learning enhanced recommendation systems are rather using deep learning as feature extractors (Hu et al., 2015a) or building end-to-end deep learning recommendation architectures (Cheng et al., 2016). However, both low level features generated by deep learning or the end-to-end deep learning recommendation system are hard to explain by IS theories. The reasons of the unexplainable RS are that the low level features which contains thousands of numerical data are hard for human to understand and the deep learning models are also highly non-interpretable Ricci et al. (2015). As such, making explainable recommendation systems for IS research seems to be an uphill task. In our research, we solve this problem by building attribute level classifier as described in Chapter 2. The attribute classifier makes the understanding of low level features becomes possible and helps us to understand RS as not a black box. In this case, we could applying IS theories into RS to explain the behavior and guide to build better RS.

To explain deep learning based RS with attribute classifier, we introduce Self Identity theory Hogg (2001) into the Recommendation Framework. In psychology, Self Identity is known as a theory to describe how social environment affects self and how self affects behaviors (Hogg, 2001; Johnson et al., 2012; Stryker & Burke, 2000). In RS, the social environment contains the product information collected from online shopping environment and the self purchase intention. How those environment factors affects self behaviors can be described by Self Identity theory. In our research, Self Identity is a set of factors attached to the self that guides behaviors in online shopping. The factors in Self Identity contains both conscious and unconscious factors (Van der Werff et al., 2013; Whitmarsh & O'Neill, 2010), in online shopping environment, conscious factors refer to the customers' clear requirements known by themselves and personal belief such as gender and culture, on the other hand, the unconscious factors include potential requirements which they do not know by themselves. An example of unconscious factors is customers may like some products by seeing it but they cannot describe which features attracted him. The previous researches are mostly focused on conscious factors (Van der Werff et al., 2013; Whitmarsh & O'Neill, 2010) because it is easy to collect. In our experiment, because of the explainable RS, we try to use Self Identity to describe the RS system and customer behaviors. In details, we split customers into five stages based on Self Identity: Harmonious, Mildly conflicted, Vulnerable, Disturbed and Fragmented (Horowitz, 2012). From Harmonious to Fragmented, people will less realize their conscious requirement and become unpredictable to their action. Therefore, the five stages refer to how well the behavior can be predicted and how clear the person knows what he wants. By Self Identity theory, we can split customers into different groups and for each group we produces different recommendation strategy.

As described above, the research contains three stages. Firstly, we build a attribute classifier to understand low level features extracted by deep learning algorithm. Secondly, we establish an deep learning based recommendation system. Finally, we apply the attribute classifier to RS to describe the shopping behavior by Self Identity theory, and based on the result, we introduced a better RS system.

5.2 Background

The research is based on deep learning enhanced RS in CS area and Self Identity theory from IS fields. To clarify our design and algorithm, we discuss the related research in both CS and IS.

5.2.1 Recommendation System

Over the last few decades, there has been a significant amount of research on Recommendation System. The Recommendation System generate suggestions to assist customers in online shopping processes, and with the help of it, customers are more likely to find suitable products (Schafer, Konstan, & Riedl, 2001). In conclude, the aim of RS is to predict how much customer like a product and ranking products by predictions (Sarwar, Karypis, Konstan, & Riedl, 2001). The former process is to evaluate how much a customer would like to buy a specific products, and in RS, it usually produces a score between 0 to 1 to describe the relationship between customer and product. The latter process is to predict Top-N products for a given customer, it will gives a list of products which the customer most preferred. Based on the algorithms and source of data, the typical Recommendation Systems are classified into three categories (Adomavicius & Tuzhilin, 2005).

Collaborative filtering RS The collaborative filtering is one the most used algorithms in RS, it build based on user-item relationships and makes the assumption that who prefer a product in the past will also like it in the future (Sarwar et al., 2001). Collaborative filtering algorithms contains a list of m user $U = u_1, u_2, ..., u_m$ and a list of n item $P = p_1, P_2, ..., P_n$. Then we can construct a $m \times n$ rating list $R_{i,j}$, for each $u_i \in U$ and $p_j \in P$ we can find the rating $r_{i,j}$ from R. Because the training data only contains part of user-item relationships, the target of training is to fill the rating matrix $R_{i,j}$. Generally, we can use Nearest Neighbors, Matrix Factorization and other Machine Learning algorithms to finish the training of collaborative filtering algorithm, and we could get all rating for each user-item pairs (Jannach et al., 2010). The CF algorithm is simple and efficient to use and predict, however, the accuracy is highly relay on the quality of existing user-item data and do not consider the features of user and item (Sarwar et al., 2001). For new products with almost no one purchased before, CF can not get accuracy ratings, and it ignores lots of product and personal information which may help to predict.

Content-based RS Content-based RS produces recommendations based on the features of items and the profiles of users (Van Meteren & Van Someren, 2000).Compared with CF, content-based algorithm recommends items that are similar to those the user liked before. It will be very efficient when predicting new items, because it only work on the description of products. However, it can hard to predict personal information because of the lack of personal data and it do not consider the community knowledge between users (Lops, De Gemmis, & Semeraro, 2011).

Hybrid RS Both CF and content-based RS have strengths and weaknesses. In this case, Hybrid RS combine CF and content-based algorithm to avoid the weaknesses for both algorithms. There are multiple ways for the combination(e.g. feature combination and feature augmentation), but it requires a large amount of data for predicting and the quality of data strongly affects the results.

5.2.2 Deep Learning based Recommendation System

As described in Chapter2, Since (Hinton & Salakhutdinov, 2006) introduced an efficient way of training deep models, deep learning has become an emerging topic in Computer Science fields. Currently, deep learning produces the state-of-the-art solutions for many problems including Recommendation System. Similar to normal recommendation system, the deep learning based algorithm can also be classified into collaborate filtering based algorithm and content-based algorithm (S. Zhang, Yao, Sun, & Tay, 2019).

Deep Collaborative Filtering Recommendation As described above, there are many algorithms can be used to predict user-item relationship $U_{i,j}$(e.g. Matrix Factorization and Nearest Neighbor), deep learning can also be used to do the training process (S. Zhang et al., 2019). Each deep learning model described in Chapter 2 can be used to do RS, and different models have different advantages. Multilayer perceptron model is used to do non-liner interaction between user preference and item features which usually performs better than liner model. Deep matrix factorization (van Baalen, 2016), on the other hand, applied deep learning to do feature extraction during matrix factorization which enrich item features in the RS. Those works use deep learning as an important part of the RS, in addition, there are also some End-to-end deep learning based models which combines the whole RS into deep learning architectures. Wide & Deep model (Cheng et al., 2016) is one of the end-to-end deep recommendation model which joint trained wide linear models with CNN to overcome the sparsity of user-item relationship. Another research trends are RNN based models which usually used to handle time series data and make user related short-term predict (Hidasi, Karatzoglou, Baltrunas, & Tikk, 2015; Y. Li, Liu, Jiang, & Zhang, 2016; Seo, Huang, Yang, & Liu, 2017). With sufficient data, deep learning based CF algorithms usually performs better than classic CF algorithms and that is mainly because deep learning models usually take more user and item features into training and predicting (S. Zhang et al., 2019).

Deep Content-based Recommendation The deep learning is also applied to content-based recommendation system. In this case, the content-based RS usually benefit for the visual features extracted from images or text features generated by Natural Language Processing (Guan, Wei, & Chen, 2019; Kumar, Khattar, Gupta, Gupta, & Varma, 2017; S. Zhang et al., 2019). Those additional features can be directly added to normal content-based RS and improve the recommendation quality. For example, in DeepStyle model (Q. Liu, Wu, & Wang, 2017), they treat visual style features along with high level category information to build style based fashion recommendation. And in the joint representation learning model proposed by Y. Zhang, Ai, Chen, and Croft (2017), they put as many as heterogeneous information sources into their model and build the state-of-the-art recommendation system.

The deep learning enhanced RS as described above usually performs better than previous shalow learning based RS, but how to explain the details of recommendation process becomes more difficult because the deep learning model is another black box for analyzing. In order to solve this problem, we use an additional attribute classifier rather than end-to-end models.

5.2.3 Recommendation and Self Identity

As reviewed in Chapter 2, previous studies on online recommendation system involve two directions: one is technical part which focus on algorithms as described above and try to increase the predict accuracy, and the other is information system part that analyzing customers' behaviors and feedbacks. Self Identity is rarely used in Recommendation System in both research trends. For CS research, they usually focus on algorithms and try to increase accuracy rather than analyzing customer behaviors (S. Zhang, Yao, & Sun, 2017), on the other hand, for IS research, the previous RS is usually considered as a black box which is hard for theoretical explanation (Van der Werff et al., 2013; Whitmarsh & O'Neill, 2010). In our research, as combined RS with attribute classifier,the RS becomes explainable, and Self Identity theory is applied to RS to distinguish customers from harmonious to fragmented.

In general, there are three major differences between the previous RS and this project:

1. In out work, we assume the interests of customers are changing over time and could be affected by outside environment. Therefore, the RS should be also changed over time for the same customer. In our project, that part is addressed by theory of self identity.

2. We considered the shopping history as a list of continuous behaviours, not just a set of single purchase. The analysis and evaluation of RS are based on a list of shopping behaviours.

3. Explainable Recommendation. We not just want to predict, but also want to figure out how the recommendation been made and what we can learn from those recommendation. In our project, instand of directly using deep features as input (Wang et al., 2017), we use human explainable mid level features as additional visual features. Based on those features, we expand self identity theory to customer shopping behavior and analyse how RS works on different kinds of customers.

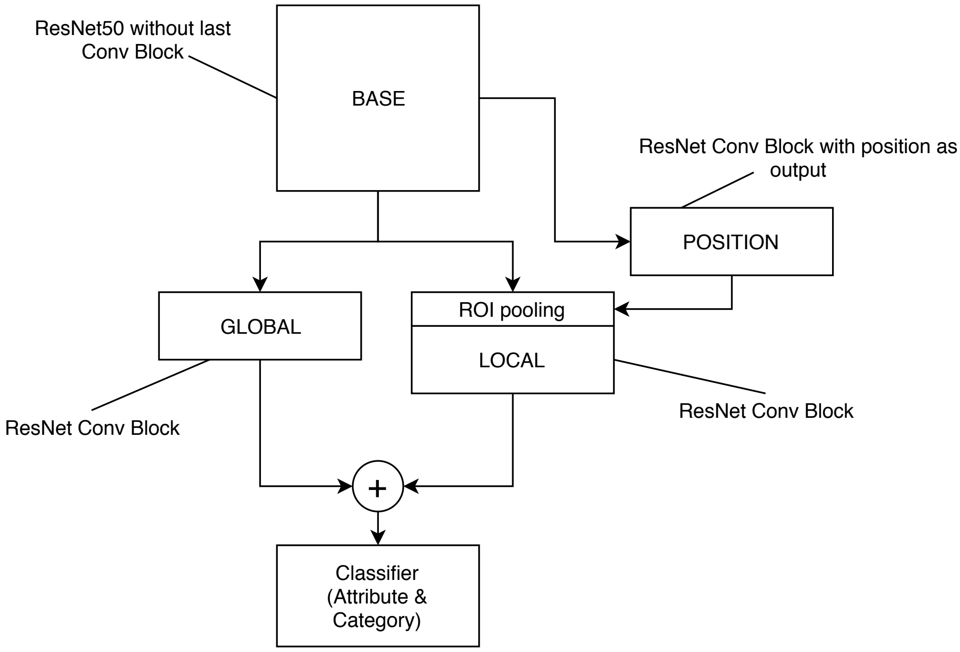


Figure 5.1: The architecture of deep learning models

5.3 Attribute classifier

In Chapter 4, we describe and implement attribute classifiers for imperfect data. Followed that research, to collect attributes from multiple datasets with multiple constraint, we present a deep learning based algorithm for attribute learning.

The datasets we used contain different types of information. In order to combine those dataset together we present a constraint transfer learning model to handle multiple attribute groups, categories and human positions. Assuming a dataset may contains three different labels of information, attributes$y_{attribute}$,categories $y_{category}$ and human positions $y_{position}$. The training model we present can handle not only all three labels together but also each single type of labels.

5.3.1 Deep Learning Architecture

The base model used in this project is ResNet50 which is relative easy and quick to train. Since the target of this project is not about accurady, ResNet is enough to produce good results for further studies. The architecture of our model is shown in Figure 5.1 include a base CNN modelM_{base}, a position detect network$M_{position}$, a global feature blocksM_{global} and a local feature blocks M_{local}. In our experiment,the M_{base} is a ResNet50 network (K. He et al., 2016) without the last CONV blocks. Both M_{global} , $M_{position}$ and M_{local} are connected to M_{base}. Meanwhile, the output of $M_{position}$ are feed into M_{local} with a special ROI polling layer together with the output of M_{base}. The final output of the model is a concatenation of the output of M_{local} andM_{global}, and there are also a special output from $M_{position}$ for position training.

5.3.2 Loss Function for Multiple Dataset

As described above, to handle different type of dataset we presents multiple models, as a result, we need multiple loss function to training the whole model. Firstly, a L_2 regression loss is used to localize the positions:

$$L_{position} = \sum_{j=1}^{N} \|v_j \cdot (\boldsymbol{\hat{p}_j} - \boldsymbol{p_j})\|_2^2 \tag{5.1}$$

where N donates number of samples,$\boldsymbol{p_j}$ refers to positions of j sample and v_j is a visibility vector, for each position in $\boldsymbol{p_j}$ 1 refers to visible and 0 refers to invisible. Secondly, two Softmax loss for position visible and categories:

$$L_{visible,i} = -f_{v_i} + \log \sum_j e^{f_j} \tag{5.2}$$

$$L_{category,i} = -f_{c_i} + \log \sum_j e^{f_j} \tag{5.3}$$

the f in equation 5.2 and 5.3 refers to the vector of class score. Thirdly, a cross-entropy loss for attributes learning:

$$\begin{gathered} L_{attribute} = \textstyle\sum_{j=1}^{N} (\boldsymbol{w}_{pos} \cdot \boldsymbol{a}_j \text{log} p(\boldsymbol{a_j}|\boldsymbol{x_j})) \\ +\boldsymbol{w}_{neg} \cdot (1 - \boldsymbol{a_j}) \text{log}(1 - p(\boldsymbol{a_j})|\boldsymbol{x_j}) \end{gathered} \tag{5.4}$$

where N is the number of samples, x_j and a_j donate the j-th images and its attributes. $w_{pos} and w_{neg}$ is the ratio of the numbers of positive and negative samples. In addition to the four loss function for four output of the models, we also introduced the triple loss Schroff, Kalenichenko, and Philbin (2015) for pairwise similar metric learning:

$$L_{triplet} = \sum_{j=1}^{N} = max(\alpha + \|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2) \tag{5.5}$$

where (x^a, x^p, x^n) is a triplet, x^a is the anchor image, x^p is the image similar to anchor and x^n is the image different from anchor. And α is the margin parameter. During training, to avoid noise data, we do not select the hardest positive and negative image as triplet. Instead, we apply the online generation strategy which random select the triplet (x^a, x^p, x^n) fit for the constraint in equation 5.6.

$$\|f(x_i^a) - f(x_i^p)\|_2^2 < \|f(x_i^a) - f(x_i^n)\|_2^2 \tag{5.6}$$

With the definition of five loss function in equation 5.1,5.2,5.3,5.4,5.5, we could handle different types of dataset in one deep learning model. In details, the choice of loss function correspond to different dataset types are shown in table 5.1.

Table 5.1: Loss function with corresponding dataset types

Datasets types	deep model	loss function
category, attribute, position, pairwise	all models	equation 5.1-5.5
category, attribute, pairwise	M_{base} M_{global}	equation 5.3,5.4,5.5
category, attribute	M_{base} M_{global}	equation 5.3,5.4

5.3.3 Training Strategy

The training of this deep learning model is slightly different because we applied five different loss function together. The strategy we used is dynamic weighted loss combination with three steps training. In the first step,the weights of loss function for $M_{position}$ is set to large, while other weight is set to small. In this stage, M_{base} and $M_{position}$ are fast trained. The second step, on the contrary, the weights of loss function for M_{global} and M_{local} are set to large. In this stage, category and attribute will be fast learned. In the final stage, the weight for equation 5.3 is set to large and the main target of attribute will be fine-turned. During training, the four stage will repeat for each batch of data until the model convergence.

5.3.4 Transfer Learning

As the present deep learning model is prepared for multiple datasets, it is necessary to define the transfer learning strategy which could involve new targets and data into models. Assuming the deep learning model is firstly trained on a datasets with all four types of label, when it is transferred to another dataset, the M_{base} as low level features will be fixed. The position model $M_{position}$ contains the position of humans which is similar in each dataset, so once it is well trained, it will be fixed in transfer leaning. The other part of the model will be trained depends on the relationship between new and old dataset D_{new}, D_{old}. For a type of category or attribute A, if $A \in D_{new} \& A \in D_{old}$ the exist global and local model will be trained, otherwise, if $A \in D_{new} \& A \notin D_{old}$ a new global and local model will be trained with the old one fixed. The rules for transfer leaning will involve new categories and attribute into model without loss the previous results.

5.4 Deep Learning based Recommendation System

As described in Chapter 5.1, to make recommendation system explainable, we do not use end-to-end deep learning RS, but applying deep features into classic RS. Based on those idea we present the deep feature based RS.

The relationship between user and item is the most import training data for RS. In our model, we consider a more complex relationships which include four different types of connections.

1. user viewed relationship. This relationship refers to a user u who viewed item X_i also viewed item X_j. In the set of relationships R,this is defined as $r_{i,j} = 1 or (0001) in binary$.

2. user view to purchase relationship. This relationship refers to a user u who viewed item X_i also bought item X_j, and it is defined as $r_{i,j} = 3 or (0011) in binary$

3. user purchased relationship. This refers to a co-purchase relationship which a user u purchased X_i also purchased X_j,and it is defined as $r_{i,j} = 7 or (0111) in binary$

4. user item purchased combination. Compared with previous relationship, this is a special time related one. It refers to a user u purchased $X_i and X_j$ the same time $r_{i,j} = 15 or (1111) in binary$

X refers to all training items and their four types of relationship belong to a relationship set R which means $r_{i,j} \in R | \forall r_{i,j} \neq 0$. The binary coded relationship means two items could have multiple relationship at the same time.

5.4.1 Recommendation System Model

There are a various models could handle those kind of dataset defined above. But to make a explainable and attribute related model, we introduce a RS with attribute space and personalize information. For every item X, we use ResNet50 K. He et al. (2016) described above to extract its deep feature x. And for each relationship described above, we try to find a distance function $d(x_i, x_j)$ to describe the relationship. For $r_{i,j} \in R$, $d(x_i, x_j)$ is smaller than $r_{i,j} \notin R$. To achieve the constraint, a suitable distance transformation should be applied. In our model, we use Mahalanobis distance Xiang, Nie, and Zhang (2008)(equation 5.7) as the base distance transformation function.

$$d_M(X_i, X_j) = (x_i - x_j) M (x_i - x_j)^T \tag{5.7}$$

Assuming F is the dimension of deep feature of X, transformation matrix M in equation 5.7 should be $F \times F$. As we use ResNet50 features, the M should contains 32 million parameters which is too large to train and can be easily overfitted. To address this issue, we follow the solution used in Der and Saul (2012) which approximate M as $M \simeq TT^T$. The matrix T is a matrix of dimension $F \times K$, and the dimension K could be a smaller number which could consider the transform as low-rank embedding.

$$d_T(X_i, X_j) = \|(x_i - x_j)T\|_2^2 \tag{5.8}$$

In addition to equation 5.8, we could consider a K dimension vector $f_i = x_i T$ for each element in X. The distance transformation could be rewrite as:

$$d_T(X_i, X_j) = \|(f_i - f_j)\|_2^2 \tag{5.9}$$

The f_i in equation 5.9 can be considered as the transferred feature space from x_i. During the training process, different feature space could be learned by selecting corresponding relationship R. As a result, we could manually control which feature we want to analyze or combine them together that makes the model more explainable.

5.4.2 Personalize RS model

The model described above only consider the relationships between user and item. The predictions are same for all users which ignore the personalize selection. Our target model need to be used by self identity theory which must contains personalize suggestion. In this case, we introduce a personalize distance function.

$$d_{T,u}(X_i, X_j) = (x_i - x_j)TD_uT^T(x_i - x_j)^T \tag{5.10}$$

where D_u is a $K \times K$ diagonal matrix, and the element in $D^u_(k, k)$ indicate how much a user u influenced by the attribute k. Similar as equation 5.9, we define a $U \times K$ matrix Y and the $D^u_{k,k} = Y_{uk}$ and then the distance could be considered as equation 5.11.

$$d_{T,u}(X_i, X_j) = \|(f_i - f_j) \cdot Y_u\|_2^2 \tag{5.11}$$

The Y_u is a personalize weight parameter which added personalized information to our model. The model we present can provide explainable attribute prediction with personalized information which can be used to analyze the RS with Self Identity theory.

5.4.3 Training strategy

For each relationship, we could use a sigmoid function transfer the distance d to probability.

$$P(r_{ij}) = \sigma_c(-d(x_i, x_j)) \tag{5.12}$$

where c is an anchor distance. If $d(x_i, x_j) = c$ then they have 0.5 probability to be related, if $d(x_i, x_j) < c$, $x_i and x_j$ have more than 0.5 probability to be related. During training, we could build a position training set with relation

R and a negative training set $N = r_{i,j}|r_{i,j} \notin R$. And optimize the likelihood based on equation 5.12.

$$\begin{aligned} L(T, c|R, N) = \textstyle\sum_{r_{i,j} \in R} log(\sigma_c(-d_T(x_i, x_j))) \\ + \textstyle\sum_{r_{i,j} \in N} log(1 - \sigma_c(-d_T(x_i, x_j))) \end{aligned} \tag{5.13}$$

The model will be learning by optimizing equation 5.13 which achieved by gradient decent. In experiment we use TensorFlow to optimize the non-linear problems.

5.5 Recommendation with Self Identity Theory

In this section, we present a novel framework combine Attribute Classifier and RS. Then we apply Self Identity theory to the framework and present the analyzing strategy and how the result can be used to improve RS.

5.5.1 Attribute Level Recommendation

From attribute classifier, we could get attribute level information for each images and based on the similarity, we could build a attribute based item-to-item relationship dataset. Assuming A_k refers to the k-th attribute, the special dataset contains $R_k = r_{i,j}|r_{i,j} \in A_k$. Applying those specific relationship dataset to RS described in section 4, we could built multiple attribute based recommendation system RS_K where K refers to a list of attribute A_k which used for training the RS. In addition, we also build personalize recommendation system based on equation 5.11, and this model will be used for analyzing personal shopping behavior with RS prediction accuracy.

In conclusion, we could produce three different RS models based on different relationship as following.

item-to-item relationship Based on item to item relationship, we build a normal deep learning based recommendation system which will be used as the base line and got global predictions.

attribute-to-attribute relationship The attribute relationship can be used to produce item-to-item relation $R_k = r_{i,j}|r_{i,j} \in A_k$ and corresponding RS can be trained. In our research, we use the attribute RS to explain how RS works and how Self Identity influence the behavior.

personalize RS Applying user-item data into RS could build personalized RS. Based on different relationship used in model, we could generate a personalize RS with selected relationship and that will results in a personalize analyzer.

The item-to-item RS and psersonalize RS have been introduced by many researchersS. Zhang et al. (2017), and the attribute RS can be regarded as a special personalize RS which use attribute relationships instead of personalize data. All of the three models have been well developed and evaluated for different situation, however, how them work in real Online Shopping Environment and what affect the different behaviors between them are rarely discussed. In this project, we will use Self Identity from Informaiton Science to distinguish different shopping behaviors and determine how different custoemrs will benefit from different RS.

5.5.2 Self Identity with View and Purchase History

To apply Self Identity to Online Shopping Environment, we need firstly demonstrate how It affect shopping behaviors and prove the relationship between RS and Self Identity. In Chatper 2, we introduce the Self Identity of a user could be considered as a combination of conscious and unconscious behaviors. For a specific user in a specific condition, different weight of conscious and unconscious factors will result in unpredictable action from Harmonious to Fragmented. In Online Shopping Environment, the harmonic behavior means the customer clearly understand what produces or features they want, on the contrary, a fragmented behavior usually result in random search and less purchases intensions. For example, the shopping process of a user search online shopping website for a red T-shirt refers to the user hold conscious requirements for red color and T-shirt category and unconscious requirement(e.g. speical pattern or texture) which may imply in subconsciousness, meanwhile, those requirements especially unconscious requirements usually changed during online shopping process. The harmonic customer tends to keep their shopping requirements during shopping, on the other hand, the fragmented customer mostly do not know what they exactly wanted.

In this Section, we firstly prove the Online Shopping behaviors can be distinguished by Self Identity theory, then based on attribute classifier, we discuss how different Self Identity could be influenced by RS. The dataset used for RS contains time serial data which can be listed to analyze a customer's behavior in a short period of time, and with the attribute prediction from RS, we could in depth analyze what attribute features the customers are concentrate on. Those attribute features in Self Identity could be regarded as conscious or unconscious factors.

The dataset designed for RS mainly focus on item-to-item relationship. In our framework, we need find a list of sequence data which a user viewed or purchased a lot of products in a short time. Based on this rules we could generate a new time series dataset for Self Identity analyzing. We set fours parameters for data generating which $time_{max}$ and $time_{min}$ which define the maximal and minimal windows size we considered for sequence data, $count_{view}$ and $count_{purchase}$ refers to the number of view and purchase behavior during the time. With the four parameters we build a sequence data scoring function

$$\begin{aligned} S_{u,t} = w_v * count_{view} + w_p * count_{purchase} + w_t * -t \\ s.t. t < time_{max}, t > time_{min} \end{aligned} \tag{5.14}$$

where u is the specific user, t is the time window size, w_v, w_p, w_t is the weights for scoring. With the scoring function, we could pick high quality data from the whole dataset. And the three weight parameters define the preference of the relationship among view, purchase and length of time.

5.5.3 Analyzing shopping behavior with Self Identity

For each sequence data we generated above, we calculate the number of item the customer viewed or purchased as $\#v, \#p$, the number of each attribute $\#A_v, \#A_p$ has been viewed or purchased, the total attribute for each image$\#a_{image}$. Then we produce a weighted attribute hit rate function$hr_i = (\#A_v^i + w * \#A_p^i)/(\#v + w * \#p) * \#a_{image}$ where w is a fixed weight to enhance the influence of purchase, i refers to the index for each attribute. We then defined the distance function d_{hr_i,hr_j} for hit rate the same as equation 5.13. The distance is scaled to 0,1 and the smaller the d_{hr_i,hr_j} is, the more dissimilar between $hr_i and hr_j$.

The hit rate function collects the distribution of attributes during online shopping, it can be considered as a attribute level behavior feature for each user. We then apply statistic methods such as average, variance to get the feature of distribution and use clustering to split user behavior to different groups. With the theory of Self Identity, we can define each group from Harmonious to Fragmented by analyzing the distribution. In details, we firstly rerank the hr from high to low which ignore the attribute information, and use K-means to cluster the reranked hr into five groups which refers to Harmonious, Mildly Conflicted, Vulnerable, Disturbed and Fragmented respectively. In cluster model, we then analyze the results in each cluster to discuss whether Self Identity can be used to describe customer behavior. In our assumption, the user with harmonious self identity tends to view or purchase products with specific attributes, therefore, the data in harmonic cluster have a larger variance and a smaller change to view more typies of attributes. On the contrary, the fragmented behaviors can be also described as smaller variance and larger view range for attributes. In detail, those feature can be described by average variance $Var(X) = \frac{1}{n}\sum_{i=1}^{|X|}(x_i - \bar{x})^2$, the number of attribute that never viewed $|A||\forall P(A) = 0$ and the number of attribute that viewed more than half chance $|A||\forall P(A) > 0.5$. In experiment, we will prove Self Identity can be used to distinguish customer behaviors by the three statistical results.

In addition, the previous research of Self Identity describes that when identity is activated, a feedback loop is also established and environmental factors will result in behaviors, meanwhile, the behavior can affect the environmental factors and finally reflect to Self Identity (Stets & Burke, 2003). The former identity behavior is analyzed by clustering the reranked hit ratio data to prove Self Identity can affect shopping behaviors, while, the latter reflection still need to be investigated. In online shopping environment, the reflection process can be described as the attribute preference changing over viewing different products. In other words, the different attributes recommended by RS will result in Self Identity behavior changing, and those changing may result in different level of pruchase intention.

Both harmonic and fragmented behaviors are easy to be explained by statistical data and those behaviors are more stable than mid behaviors. To analyze the reflection process of Self Identity, we choose the mid behavior between harmonic and fragmented as the datasource. In details, we select mid behaviors with longer duration and split their shopping behavior with small sliding window to determine the attribute based Self Identity changing over shopping. There are three different kinds of data can be found as following.

From harmonious to fragment The three variables, average variance, number of never viewed attributes and number of usually viewed attribute, change from harmonious level to fragment level which means the customer firstly concentrate on specific items, but changed to uncertain products during Shopping. This process can be described by two different situations. The fisrt one is a customer prefer to specific attribute features but be attracted by other attributes during shopping. The second one is a customer focus on a specific products and purchase them, then he is attracted by other different products.

From fragment to harmonious The three variables described above change from fragment level to harmonious level. This process describes the customer find their preferred attribute combination during shopping and usually result in high purchase intension.

Keep mildly We can not find statistical relationship from the variables described above. That may because the customer keep changing perference during shopping or the attributes we used are not enough to describe this kind of user.

As harmonious behavior usually result in high purchase intention, to develop different strategies for different mid behaivor customer is necessary. Specially, for users chaging from fragment to harmonious, we prefer to enrich the attribute range predicted by RS to accurate this process and lead to higher purchase rate.

In experiment, to get longer analyzing data, we increase the time weight w_t in score function 5.14 and get high score data. Then we apply sliding window to generate time series data. Finally, we use the Self Identity cluster model to analyse the changing of behaviors for mid behavior custoemr.

5.5.4 Enhanced RS with Self Identity and Attributes

Based on the research on how Self Identity affect shopping behavior, accuracy is not enough to describe the quality of RS, especially considering the purchase intention. For each user we could generate a specific hit rate function during shopping, those hit rate refers to how a customer concentrate on specific attributes. As we trained multiple attribute relationship RS, we present a combined recommendation system with Self Identity and attribute attention.

$$\begin{aligned} P = (1-\alpha)P_{T,u} + (\alpha-\beta)\tfrac{1}{|A|}\textstyle\sum_s^{|A|} hr_s \cdot P_s \\ +(\alpha+\beta)\tfrac{1}{|A|}\textstyle\sum_s^{|A|} hr_s^{-1} \cdot P_s \end{aligned} \tag{5.15}$$

where α is the Self Identity factor, it will be larger for fragmented user and smaller for harmonious user. β is the attribute concentration factor, the model will focus on specific attributes when the value is small. $P_{T,u}$ refers to the normal RS and P_A refers to a list of attribute RS. The combined model takes advantage of the attribute relationship and uses Self Identity to describe how user concentrate on specific attributes in a short time to make better predictions. Meanwhile, for Mid behavior users, we try to increase the prediction variety rather than the accuracy which may encourage their behaviors changing from fragmented to harmonious which may increate their purchase intention.

5.6 Experiments and Analysis

As described above, the experiment include three stages, for every stage we use different data and generate different models. The first experiment is to train a attribute classifier by using transfer learning from three datasets. We then get a complex attribute classifier for hundreds attribute. The second target is to build deep learning enhanced Recommendation System. Following the protocol described in section 5.4.3, we trained nine different models for both global and attribute recommendation. Based on the above two experiments, we applied Self Identity to the RS and shopping behavior to discuss how Self Identity can help to identity shopping behaviors and improve RS. Here we literary discuss the experiment for each stage.

5.6.1 Attribute Classifier

5.6.1.1 Dataset collection

For attribute classifier, we use transfer learning strategy and train the model in three different dataset. The dataset we used is described in Table 5.2. The DeepFashion dataset (Z. Liu et al., 2016) contains most of information including category, attribute and human position, but the variety of attribute is still limited. To solve this problem, we select other two datasets which contain more different semantic attribute and learn those different attributes to one model.

Table 5.2: Datasets for Attribute Learning

Datasets	# image	# attribute	# category
DeepFashion (Z. Liu et al., 2016)	181,072	85	12
CwA (Farhadi et al., 2009)	2,171	36	5
Fashion10K (Loni et al., 2014)	18,487	31	2

5.6.1.2 Training Strategy

We firstly train the whole model include M_{base},M_{local},M_{gloabl},$M_{position}$ on DeepFashion dataset which include all the four typies of information described in Section 5.3. The deep learning model use Adagrad algorithm (Kingma & Ba, 2014) to do back propagation. And as suggested in Section 5.3.3, the training weights for different loss functions are 0.90 for position learning and 0.1 for feature learning at the beginning, and be slightly changed to 0.05, 0.95 for position and feature learning at the end. Next, we apply transfer learning to train the model on CwA and Fashion10K datasets (Farhadi et al., 2009; Loni et al., 2014) in which not all typies of information are included but enrich the attribute range. During training, we fixed M_{base}, M_{local} and $M_{position}$ models and trained another two M_{global} models for each new dataset.

The deep learning model was implemented in Python with TensorFlow framework (Abadi et al., 2016), and the training was applied on two AWS G2 instances (Amazon, 2015). It totally cost 17 hours to train the model for all the three datasets. After training, we generated six models: one M_{base}, $M_{position}$ and M_{local} model trained on DeepFashion dataset and three M_{gloabl} models trained on all three datasets.

5.6.1.3 Results

The attribute classification model is trained based on the three different dataset we discussed above with total 200k images. To evaluate the accuracy of model, we keep 80%,10%,10% as train,evaluate and test samples separately. The attribute classifier we trained could classify 113 different binary attributes, but not all attributes could be clearly classified. We removed 31 attributes which the model can hardly predict, and 13 attributes which performs quite different among different datasets. Then we group the rest attributes into 9 groups by attribute grouping (e.g. color related attribute will be moved in color group) and list corresponding accuracy as described in table 5.3. For baseline model, we select exact the same settings and use origin Resnet50 architecture for comparing.

Table 5.3: Accuracy of the Attribute Classifier

Attribute groups	Colors	Patterns	sleeve	Neckline	Texture	Fabric	Shape	Part/Category	Style	Total
# Attribute	15	7	3	3	9	6	8	10	8	69
Accuracy of Baseline	0.773	0.6919	0.9127	0.9342	0.6577	0.6136	0.5249	0.7576	0.7004	0.7296
Accuracy of Ours	0.8972	0.723	0.9131	0.9421	0.7421	0.6543	0.6132	0.7543	0.7143	0.7726

The accuracy is not the important part of our project, but the result still prove our proposed attribute learning architecture outperform the origin one especially for Color, Pattern and Shape attribute groups. That is mainly because our model can take use of the position information and pairwise constraint to avoid the influence from background noise. To illustrate the result more clearly, we select three samples which is believed to be the good, average and poor prediction results shown in Figure 5.2. The red underlined attribute refers to the incorrect prediction and the attribute in brackets is the ground trues.

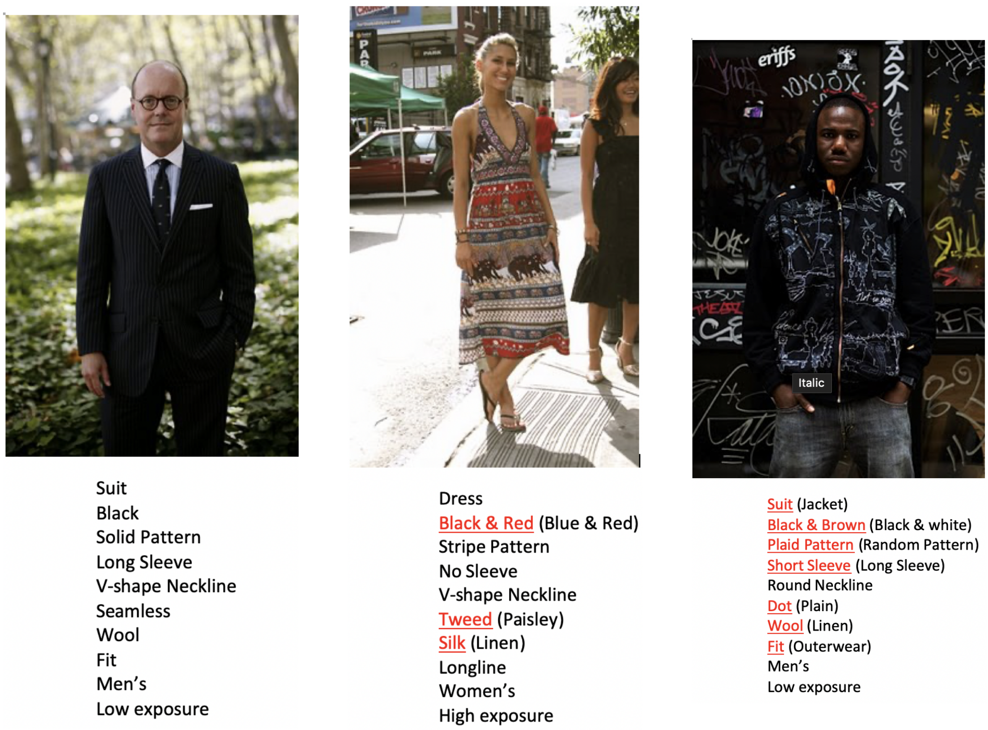


Figure 5.2: Example of good, average and poor prediction results

5.6.2 Recommendation System with deep features

5.6.2.1 Dataset collection

We use Amazon clothing item-to-item dataset (K. He et al., 2016) as the base dataset. The details of the dataset are listed in Table 5.4. The dataset contains four different relationships as described in Section 5.4 and we consider those as the dataset-global for training global RS. There are too many users($>$1M) in dataset, to build personalize RS, we only select active users(view or purchase behavior > 30) with related relationship as dataset-personal for training personalize RS. For attribute related RS, we firstly collect product's images and identify the quality of them. The quality of image is defined as the visualability of human in image which can be determined by $M_position$ model from Attribute classifier. In Amazon Fashion dataset more than 99% items contains high quality image. We then apply attribute classifier generated above to products' images to get attribute labels and then build a attribute relationship dataset as dataset-attribute. In experiment, we build 8 attribute RS based on dataset-attribute.

Table 5.4: Amazon Clothing Dataset Settings

Dataset	#User	#item	#rating	#relationship
Amazon Fashion Dataset	3,260,278	773,465	25,361,968	16,508,162
Dataset with Qualified Image	3,031,260	741,019	22,019,256	15,972,145
Dataset for Personalize RS	1,985,463	694,184	19,236,483	12,693,360

5.6.2.2 Training Strategy

Considering the efficiency and accuracy, we choose different sizes of K as the transform rank, the 1000 dimension output from last ResNet50 layer with AVG pooling as the image feature. As discussed in Section 5.4.3, we trained 9 different models: one item-to-item relationship model, 8 attribute relationship model and 1 personalize model. For each model, it takes 4 hours to coverage. The training protocol is described below:

1. To find the best parameter for each RS, we choose multiple K values from 10 to 100 and evaluate which is the best choice.

2. We collect the R data as positive set and random select equal size of unrelated data as negative set. Then we split those data into Train, Validation, Test dataset with $\%80$,$\%10$,$\%10$ data respectively.

3. Training model with parameters generated above on the dataset and evaluate the results with the assumption that $P(x_i, x_j) > 0.5 | \forall r_{i,j} \in R$.

4. Select the best K value for each model and report the evaluation result on Test dataset.

5.6.2.3 Results

The accuracy of each model is described in Table 5.5. The accuracy is calculated based on the protocol 3 described above in section 5.6.2.2. For any two randomly selected products, the proposed model should make a prediction bigger than 0.5 if they contain some relationships during shopping, otherwise it means the model fails to do the recommendation. In the experiment, all the 9 models reach considerable accuracy which is enough for further Self Identity analyzing.

Table 5.5: Accuracy of Recommendation Models

Model	Parameters	Accuracy
Global-RS	K=100	0.897
Personlize-RS	K=50	0.9038
Attribute-RS-Color	K=10	0.8073
Attribute-RS-Patterns	K=25	0.8654
Attribute-RS-Sleeve	K=10	0.911
Attribute-RS-Neckline	K=10	0.9272
Attribute-RS-Texture	K=25	0.8386
Attribute-RS-Fabric	K=25	0.8175
Attribute-RS-Shape	K=25	0.9321
Attribute-RS-Part	K=25	0.9143
Attribute-RS-Style	K=25	0.8014

5.6.3 Analyzing Shopping Behavior with Self Identity

The personalize recommendation system do not split user into different groups but only consider the user-to-item relationship. As a result, it is hard to predict the behavior for users with less relationship, then the personalize recommendation system is more close to global recommendation system. In this experiment, we predict preferred products for users in time series dataset and apply hit rate function to analyze as described in Section 5.5. The distance between different hit rate hr is defined similar to equation 5.12 in which the distance is transferred to probability from 0 to 1.

5.6.3.1 Dataset collection

As described in Section 5.5.3, we select time serial shopping data from Amazon Clothing Dataset (K. He et al., 2016) and use the scoring function 5.14 to determine the quality of data. In experiment, to get meaningful data we set the $time_{max} = 4$, $time_{min} = 0.5$ in hours and $w_v = 0.1$, $w_p = 0.5$,$w_t = 2$ for weights. The score describes whether a user viewed more than 10 times or purchased more than 2 times in half hours. The equation part $w_t * -t$ will make sure the length of sliding windows to be shorter and the frequency of behavior to be larger. The table 5.6 illustrates the data average score before and after scoring filtering with $score > 0$. The result proves our data collection strategy change the dataset from item-to-item relationship to user behavior focused. The dataset will be used to do Self Identity related research.

Table 5.6: Dataset properties with score filtering

	#Record	#User	Avg. Score	Avg. Time(hours)	Avg.View	Avg.Purchase
Total Records	1260278	1260278	-1.1757	0.59	0.033	0.002
Selected Records	31341	30172	1.315	1.01	31	0.47

The dataset for mid behavior analysis, shown in Table 5.7, is generated based on the above user behavior dataset. In order to get long history data, we set $w_t = 4$ and $w_v = 0.05$ to rescoring the data and remove the data with $score < 0$. For the sliding window splitting, we set the window size to 0.2 hours and move step to 0.1 hours which we can get 9 sliding windows for 1 hour data. This dataset will be used to analyze mid behavior users.

Table 5.7: Dataset for analyzing mid behavior users

	# Data	Avg. Score	Avg. Time(hours)	Avg.View	Avg.Purchase
Total Record	31341	1.315	1.01	31	0.47
Selected Record	3517	0.2255	1.35	56	0.51

5.6.3.2 Customer Behavior Clustering with Self Identity Theory

The approach of customer behavior clustering is used to prove how Self Identity influence the shopping behavior. Following the strategy proposed in Section 5.5.3, we introduce the following experiment protocol.

Quality check and feature extraction we firstly collect image data from the dataset collected above and justify the quality by using the visualability predcition from $M_{position}$ model. The low quality image will be removed, and the others' feature will be collected from ResNet50 model.

Attribute feature prediction The attribute classify model predicts attribute features for each image and the result will be stored in a vector.

Clustring with reranked hit rate With proposed method in Section 5.5.3, we calculate the reranked hit rate histogram and its statistic features for each user, and then group all users together to do clustering with K-means. Finally, we report the clustering result with associated models.

5.6.3.3 The influence of Self Identity in Online Shopping

We cluster the data into five groups based on the hit rate histogram and statistic features shown in table 5.8.

Table 5.8: Clustering result

Clustering groups	# User	Avg.Variance	Avg. P(Attribute)>0	Avg. P(Attribute)>0.5	Avg.#Purchase
Group1	2437	0.039665133	47.3712	6.1821	1.7207
Group2	4841	0.029625402	51.321	3.9812	0.7721
Group3	10394	0.02603529	51.9742	3.1215	0.5321
Group4	1949	0.024057834	52.8431	2.9736	0.2415
Group5	11720	0.019460588	61.1732	1.2643	0.0702

The Avg. Variance refers to the average hit rate variance for each customers, the higher variance means the customer concentrate on particular attributes, on the contrary, lower variance refers to the user prefers to do random search products. The Avg. P(Attribute)>0 indicate the average number of attribute the customer viewed during online shopping, similarly, Avg. P(Attribute)>0.5 refers to the average number of attribute that the customer viewed more than 50% chance.

From the result, we could clearly find the customer in Group1 tend to behave more harmonious and concentrate on special attributes. And for customers in Group 5, they behave fragmented and can not be distinguished by attributes' differences. The rest of users in group 2,3 and 4 perform similarly that the behavior of them is between group 1 and group 5. Based on the definition of Self Identity, the users in group 1 can be regard as Harmonious, the users in group 5 is Fragmented and the rest user is the Mid behavior customer. In detail, we random pick three customers from group 1, group 2,3,4 and group 5 respectively to illustrate the different histograms among different groups.

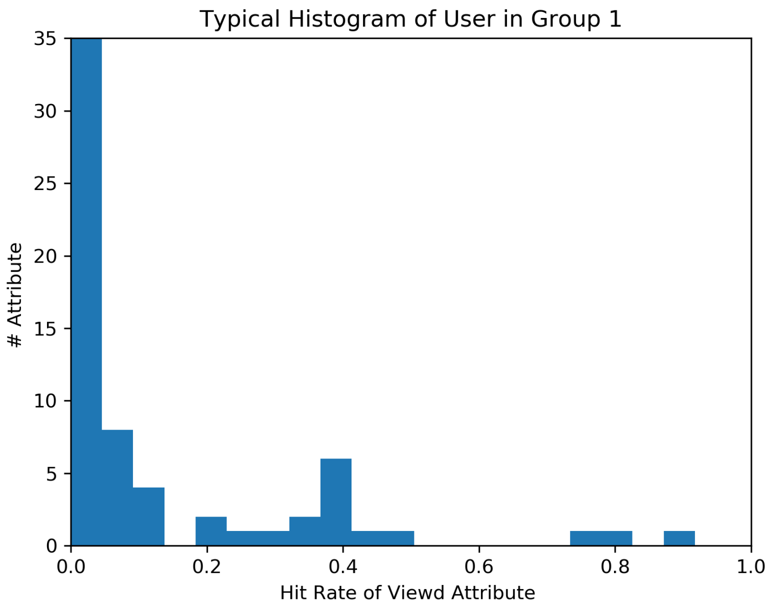


Figure 5.3: Typical histogram of user in group 1

The figure 5.3,5.4,5.5 illustrate three typical histogram of reranked hit rate in three different groups. It is clear that a customer with harmonious identity tends to focus specific attributes and ignore a lot irrelevant attributes. On the contrary, the user with fragmented identity randomly viewed every attributes. And the users in group 2,3,4 concentrate on some attributes but are not very convinced.

The result also indicate that customer with harmonious purpose usually result in higher purchase chance. For Recommendation System, to encour-

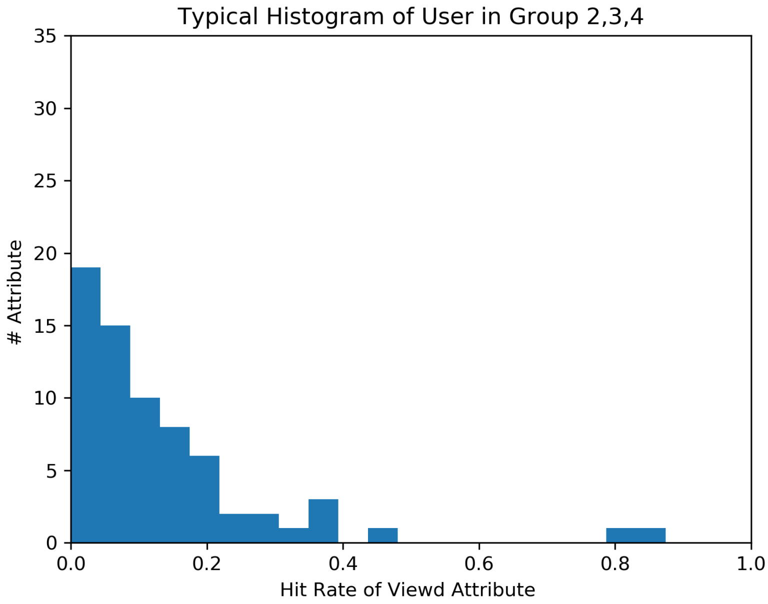


Figure 5.4: Typical histogram of user in group 2,3,4

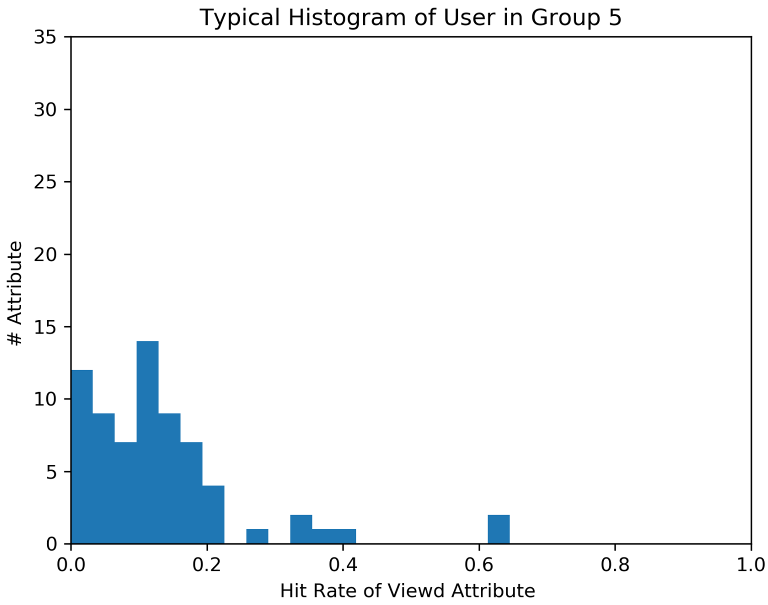


Figure 5.5: Typical histogram of user in group 5

age online shopping customer changing from fragmented to harmonious is important than just increase prediction accuracy.

5.6.3.4 The reflection from RS to Self Identity

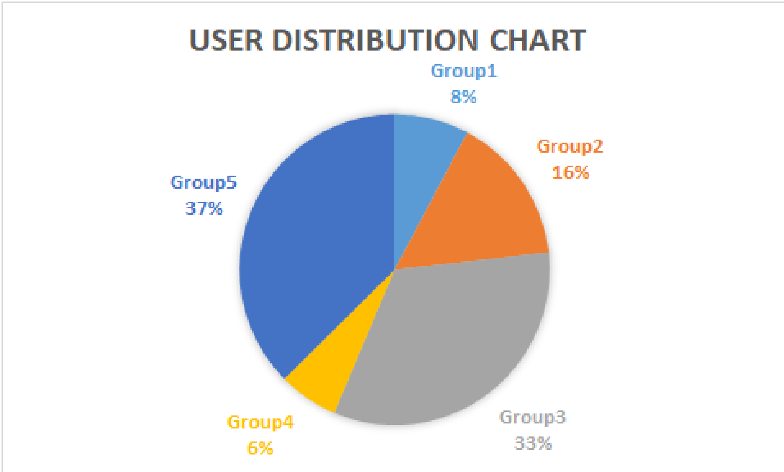


Figure 5.6: User Distribution Chart

As shown in figure 5.6 more than half customers are Mid behavior customers, it is important to understand how to build better RS for those users. As Self Identity can be affected by the reflection from environment, in section 5.5.3 we also propose a sliding window based method to describe how the Mid customer behavior changed over time. We calculate the variance, the number of attribute never viewed and the number of attribute viewed more than half chance in each sliding window, and the differences between each window. If the variance increase more than 0.001 for a period of time the customer marked as Fragmented to harmonious, on the contrary, the variance decrease more than 0.001 will be marked as harmonious to fragmented. The rest customers are labeled as Mid user. The result is shown in Table 5.9.

Table 5.9: Clustering result

Type	# User	Avg.#Purchase	Avg.Variance	Diff. Variance	Diff.P(A)>0.5
F to H	3792	0.9131	0.029936	0.002139	1.134
H to F	6572	0.374	0.025431	-0.002732	-0.294
Mid	6820	0.559921378	0.2329	-0.0001021	0.072

The three types are discussed in Section 5.5.3, and the Fragmented to Harmonious customer more tend to purchase products compared with Harmonious to Fragmented one. The results indicate that the Self Identity of users could change over time, and the changing may because they meet the attribute combination they preferred. In this case, increasing attribute combination range of recommendation may promote the Self Identity of customer to change from Fragmented to Harmonious.

5.6.4 Improve RS with Self Identity

The results above show the Self Identity theory is necessary for analyzing shopping behaviors, especially the behavior happens in a short period of time. The user with harmonious belief in products will concentrate on specific attribute combinations. Thus, applying attribute based recommendation model may help to increase the prediction accuracy for that group of customers. In this experiment, we implement the algorithm proposed in Section 5.5.4, to determine whether Self Identity classification can be used to improve current RS. Meanwhile, for analyzing Mid behavior customers, we select the sliding window dataset and for each user we classify their behavior changing by clustering model generated above. Based on the result, we will choose different recommendation strategies for them.

5.6.4.1 Dataset Collection

We use the same data described in Section, and to split them into Train, Validation, Test dataset.

5.6.4.2 Results

As discussed in Section 5.5.4, parameters α and β are used to control the joint model. In our experiment, we use n-fold to select best parameters which do not only consider the recommendation accuracy but also consider the attribute recommendation range. The α is set to 0.95, 0.8, 0.8 and β is set to 0, 0.1, 0.05 for Harmonious, Mid and Fragmented customer respectively.

Table 5.10: Accuracy of Recommendation System with Self Identity

Model	Accuracy for All	Accuracy for G1	Accuracy for G2,3,4	Accuracy for G5
Personlize-RS	0.9038	0.9172	0.9091	0.8716
RS with Self Identity	0.9013	0.9318	0.9087	0.866

The Table 5.10 shows the result of different RS for different group of users. It clearly points out that the prediction for fragmented users is difficult than others because they are more prefer to do random view. By taking use of the attribute information, the prediction accuracy for harmonious users is increased which indicate the Self Identity can be used to describe the online shopping behavior.

Table 5.11: Average number of attribute in Top 10 predictions of Recommendation System with Self Identity

Model	Avg.#Attribute	Avg.#Attribute for G1	Avg.#Attribute for G2,3,4	Avg.#Attribute for G5
Personlize-RS	46.12	41.13	44.96	47.04
RS with Self Identity	49.92	37.25	51.86	49.75

Another scope of evaluation, as discussed above, is the average number of attribute in top 10 predictions of RS. The table 5.11 describes the result of how Self Identity works on RS. It proves that the RS with Self Identity will more concentrate on specific attributes for Harmonic customers, meanwhile, it increase the predicted attribute range for Mid and Fragmented users. The new attribute combination may push mid customers to become harmonious and result in better purchase intention and shopping experience.

5.7 Summary

In this experiment, we propose a framework that combine the deep feature based recommendation system with Self Identity theory. The contribute of this work is in two-fold. On the one hand we prove the Self Identity can be used to describe shopping behavior affected by different attributes they viewed, on the other hand, we propose the new recommendation strategy based on Self Identity theory. The results of experiment also indicate the accuracy from Computer Science is not enough to describe the quality of RS, for Mid customers, enrich the attribute combination is more important than predicting related products.

Chapter 6

Discussion, Conclusion and Future work

The reviews and exeperiments describe the research gap between Online Shopping Experience in Information System and Machine Learning techniques, especially Computer Vision from Computer Sicence. And we proposed two methods to solve the lack of data problem and combine the Informaiton System theory into Machine Learning algorithms. In this chapter, we will discuss and explain the research questions and the experiment we used to answer the questions.

6.1 Discussion

6.1.1 Research Gap between Machine Learning and Online Shopping Experience

The thesis, improving Online Shopping Experience by Knowledage discovery from communities, is a big topic include two popular research fields: Online Shopping Experience in Information System and Knowledge discovery from Computer Science. Both research areas are also include huge research topics. To clearify the connection between Online Shopping Experience and knowledge discovery and propose our research targets is our first work in the thesis.

The target of research in Online Shopping Experience is to define and discuss what factors in online shopping procedure can affect Customers' shopping experience. In the review, we describe that the early researches in Information System area are concentrate on high level overview of Online Shopping Experience which regard the Online Shopping Environment as a whole system or only spilt them into 4 or 5 parts (Frow & Payne, 2007; Grewal et al., 2009; Häubl & Trifts, 2000; Meyer & Schwager, 2007). With the developing of online shopping, the components of Online Shopping Environment has been treated as more detail particulars. To evaluate Online Shopping Experience, we can analyze the problem from a lots of subset of Online Shopping Environment such as Web Service (Kaynama & Black, 2000), Web Design (Dix, 2009; Zviran et al., 2006), Online Customer Behavior (Moorman et al., 1992; Morgan & Hunt, 1994; Rousseau et al., 1998) and Online Recommendation System (Childers et al., 2002; Park & Kim, 2003). In this case, improving Online Shopping Experience can be regarded as a problem of improving a subfields listed above. By reviewing all those subfields, we conclude that Online Recommendation System is the most related field which can be improved by combining Information System theories and Knowledge discovery methods.

There are three reasons to focus on Online Recommendation System. Firstly, Online Shopping Recommendation System has been developed in both Information System and Computer Science area, but the research target is different. The research in Information System tries to analyze human response to the RS and how RS helps in Online Shopping Environment (Childers et al., 2002; Park & Kim, 2003). On the other hand, researchers in Computer Science mostly focus on the performance and how new elements from websites can be used in RS (H. Chen et al., 2012a; Di et al., 2013a). Secondly, our research methods, knowledge discovery from communities, need huge data to analyze and produce results. Coincidentally, the online shopping recommendation system is built based on those large amount of data which are a good resource for our knowledge discovery (Hu et al., 2015a). Finally, there is a research gap in Online Recommendation System between Information System and Computer Science. The RS in Information System is usually treated as a black box, researchers do not want to consider how RS works and what the prediction procedure means in Online Shopping Experience. The researches in Computer Science area, on the contrary, are all data centered that they only interest in mathematics or algorithms in RS, the feedback from customers is rarely considered (G.-G. Lee & Lin, 2005). In conclusion, to fill the research gap which investigate RS as a white box and improve RS with the results from Information System becomes our target.

Knowledge discovery, as our research methods, is another big topic in Computer Science. The sub-fields of knowledge discovery include Data Mining, Big Data analysis, Information Retrial and Machine Learning. In Chapter 2, we discussed the recent research trend (Lecun et al., 2015) in Machine Learning and how Machine Learning helps RS Iwata et al. (2011). We particular focus on Computer Vision, a sub-field of Machine Learning, to discuss how it helps RS in Information System view. The reason to choose Computer Vision is in two-fold. On the one hand, Computer Vision attracted considerable attentions due to the developing of Deep Learning and it provides the ability to analyzing complex image data which can help to understand the RS (Lecun et al., 2015). The Computer Vision provides the potential to split Online Shopping Experience into more detail parts. On the other hand, RS is built based on a large amount of data which is a good resource for Computer Vision. Actually, there are already a lot of research in improving RS with image features, but most those research are built on well picked data rather than real online shopping environment (H. Chen et al., 2012a; Di et al., 2013a). Therefore, how Computer Vision can help Online Shopping Experience, especially Online Recommendation System, and how it work in real Online Shopping Environment becomes another research target.

Here we propose our research target in both Computer Science and Information System field:

Computer Science The Computer Vision methods are mostly developed on well picked data, but real Online Shopping Environment contains lots of noisy and unlabeled data. So the weekly supervised learning problem of extending current Computer Vision methods to handle noisy and partially labeled data in online shopping environment is our first research target.

Information Science Most research in Information System regards RS as a black box which do not analyze the inside the recommend approach. The second target is to introduce a theory to describe RS behavior and improve RS with Information System theories.

Those two research targets are introduced based on reviewing related work and try to fill the research gap between Online Shopping Experience and Computer Vision. The result of our research could improve Online Shopping Experience in both Computer Science and Information System fields.

6.1.2 Machine Learning in Online Shopping Environment

To solve the first problem, we firstly describe a real online shopping scenario. Semantic attributes, compared with high level categories, contains more information about images in Online Shopping Environment Hu et al. (2015b). For example, the high level category for clothing images only split clothes into limited groups that a dress with different style can not be distinguished by the high level features. Semantic attributes for clothes contain more information such as style, color, texture and so on. Those additional information can meticulous recognize images which is necessary for Information System research. However, the complex classification requires high quality well-labeled data which is hard to collect in real Online Shopping Environment. Therefore, we present an algorithm to improve attribute classification with imperfect training data H. Chen et al. (2012b); S. Liu et al. (2012).

The target of the problem is training semantic attributes classifier, the data is a small set of well-labeled data which is not enough for supervise training and a large set of noisy pair-wise data. The setting of data is easy to find in real Online Shopping Environment, the large set of noisy pair-wise data can be generated by purchasing relationship or collecting from inexpert annotator, the small set of accurate data can be get from expert annotator. This is a much easier and cheaper way to get semantic attribute training data compered to previous methods with fully labeled data. As described in Chapter 4, we present PCRF algorithm to classify and reduce noisy pair-wise data and change them to labeled data, then the data can be used by normal algorithms. The results show our algorithm outperforms previous similar methods and reach similar accuracy with well-labeled data.

In conclusion, for the first target, we propose a general Online Shopping scenario and solve the problem by weakly supervised learning. The experiment results in two findings, firstly, it is necessary to create new algorithms for Online Shopping Environment which include more complex and noisy data. The present Computer Vision algorithm may fail to handle imperfect data or can not get reasonable results. Secondly, with the developing of Computer Vision, it is possible to create new strategies to handle complex data. In recent years, the deep learning significantly increases the predict accuracy into vary high level, in some special case, it already outperforms human. We could get various accuracy prediction with new models and those result may provide new ways for Information System research.

6.1.3 Analyzing and improving Recommendation System with Self Identity theory

This experiment concentrates on the second problem, how RS can help to analyzing Information System theory and how the theory improve RS. To analyzing RS, we firstly try to make RS explainable which is easily to be done in mathematics but hard to be analyzed in human behavior aspect (Ricci et al., 2015). For Information System research, to explain human behavior predicted by RS is important. We explain that by adding more meaningful features rather than latent variable to current RS. The meaningful feature we used is the semantic attributes from images. The first experiment answers a question that how to build attribute classifier from imperfect relation data. The dataset for RS can be regarded as noisy relation data which can be used by the algorithm proposed above and then attribute for images can be extracted. We then present deep feature based RS and semantic attribute based RS. As described in Chapter 5, those RS can be explained by analyzing attribute features from images.

The results of the explainable RS can be described by Self Identity theory which split shopping behaviors from harmonious to fragmented (Horowitz, 2012). A person with harmonious identity tends to purchase products with particular features, on the other hand, a fragmented customer usually random search items. With the theory, the analyzing results prove that harmonious customers can be easily predicted, while, fragmented behaviors are predicted in low accuracy. Those human behaviors and RS results are integrated in Self Identity theory in which the RS can be explained by Information System. Based on the result, we present a multiple attribute related RS for harmonious customer which outperforms the current one. But to make RS predict fragmented behavior, the model is easy overfitted and can not give meaningful prediction.

Rather than the accuracy of RS, we try to describe quality of model with different ways. We analyzing the mid behavior between fragmented and harmonious and found those behavior can be meticulous spitted to two stages. The first stage is random searching which can be regarded as fragmented behavior, and the second stage is more close to harmonious which they found some particular attribute combination they want. The progress to encourage the Self Identity from fragmented to harmonious can not improve RS accuracy, but it is necessary in Information System theory which improve the Online Shopping Experience. In this case, we presents a joint RS which for harmonious customers we try to make high accuracy prediction and for fragmented customers we try to increase the range of prediction.

The results of this experiment describe that how Information System theory can help in RS. The targets of research, compared with pure Computer Science problem, are not all about accuracy but also include behavior analyzing which makes the model more suitable for different kinds of customers, even it is not increase the accuracy. It also point out the differences between Computer Science and Information System, but with suitable theory, the differences can also be used to improve online shopping experience.

6.2 Conclusion

This thesis firstly reviewed recent works in Online Shopping Experience and discussed how Knowledge Discovery method such as Machine Learning and Deep Learning can help to improve Online Shopping Experience in Chapter 1 and 2. Based on the review, we propose two research targets, one is Machine Learning in real Online Shopping Environment and the other is how Information Science theory can help to improve Machine Learning techniques. In Chapter 3, 4 and 5, we designed two experiment to solve the problem addressed above and got reasonable result which prove the potential of Machine Learning to be used in Information Science research. Finally, we discussed the reason to choose those two experiments and the inspires we got from the result.

Online Shopping Experience and knowledge discovery are two big topics in both Information Science and Computer Science, it is impossible to solve the problem in all aspects within one research. In this case, we particularly focused on sub-fields of those research which is Recommendation System from Online Shopping Experience and Machine Learning from Knowledge Discovery. In detail, we concentrate on two problems which are Machine Learning in Online Shopping Environment and improving Recommendation System with Self Identity Theory. The contribution of this thesis is in three aspects.

Research Gap between Online Shopping Experience and Machine Learning
we reviewed recent research on both sides and conclude the research gap between them. In Online Shopping Experience, the research is

mostly focused on existing design and algorithm, how new methodologies like Machine Learning can help is rarely discussed. Meanwhile, it usually regards the algorithm or system in Online Shopping environment as a whole black box that does not analyze the in depth details under the algorithm or system. In Machine Learning, the algorithms are mostly built on well picked datasets which are hard to get in Online Shopping Environment, how to apply Machine Learning to real Online Shopping Environment is still a problem.

Attribute Learning with weakly supervised data As described in Chapter 4, a possible way to combine Machine Learning with Online Shopping Environment is to built a semantic attributes classifier based on the realistic data collected from online shopping website. Therefore, we propose a Pairwise Constraint Random Forest to do classification and noise data removal. The results prove the algorithm we proposed can work with imperfect data and the potential of applying Machine Learning algorithm to more complex scenario.

Recommendation System with Self Identity theory In this project, we implement deep feature based Recommendation System and prove the customer behavior can be distinguished with Self Identity theory. For different customer groups, we point out that different Recommendation targets are needed for different group of customers, and then we propose a joint Recommendation System which do not only consider the accuracy but also consider the Self Identity behaviors. This research also demonstrate that in real online shopping environment, statistics evaluation methods from Computer Science are not enough to describe customer behaviors. Some theories from Information Science are necessary to improve Online Shopping Experience.

The three contributions prove the novelty of our research, but there are still limitations existing in our work. And based on those limitations, we can conclude the implications in further research.

6.3 Limitations and Implications

As discussed above, Online Shopping Experience and knowledge discovery contain lots of subfields, and our research can not cover all aspects of those two fields. In addition, for each specific scenario(e.g. weakly supervised data in an online shopping environment), there is still a lot of different situations. For example, the weakly supervised data can be pair-wise or group-wise, and the noisy can exist in image level or label level. In theoretical development, we prove Self Identity can be used to enhance Recommendation System, but we can not confirm Self Identity theory itself is sufficient for describing the Recommendation System. We may need to introduce more theory from Information Science to describe the work of Recommendation System. In detail, we list the limitations and implications of this thesis as following:

Complex weakly supervised data To apply machine learning to real online shopping environment, we need to handle the complex and imperfect data. In our research, we consider the special condition that we already got a small set well labelled data and need to use another large set pair-wise imperfect data to increase the prediction accuracy. However, the imperfect data could be in another situation. Firstly, the noise of data could be in both image-level and label level, which means the image could be unclean and mislabeled at the same time (Mnih & Hinton, 2012). In our work, we only consider the label level noise and ignore the unclean images. Secondly, the imperfect data can be more complex group-wised rather than pair-wise (Claesen, De Smet, Suykens, & De Moor, 2015). The group-wise data refer to the attribute of some data in one group may differ from that in other groups which describe a more general situation but more difficult to handle. Finally, the well labelled data can not exist or quite limited, which can be regarded as zero-shot (Socher, Ganjoo, Manning, & Ng, 2013) or one-shot learning (Fei-Fei, Fergus, & Perona, 2006).

Complex prediction target In online shopping environment, the category level prediction is not enough for improving the online shopping experience. Thus, we choose semantic attribute learning as the target and use the result to analyze and improve the recommendation system. In recent years, with the development of deep learning, more complex prediction target can be used in the online shopping environment. For example, the sentiment analysis from customers' feedback (Kouloumpis, Wilson, & Moore, 2011) which can be used to understand more personalize feelings from comments and help to describe the customer more clearly. The generative adversarial network (Goodfellow et al., 2014) which generate new items rather than classify them can be used to generate new online shopping items based on personal preference, and it will help for the company to develop new products (S. Zhu, Urtasun, Fidler, Lin, & Change Loy, 2017).

Others fields in Online Shopping Experience As discussed in Chapter 2, the Online Shopping Experience contains a large number of research fields. In our project, we particularly focus on the Recommendation System as it has been discussed a lot in both Information Science and Computer Science. There are still a lot of other fields in Online Shopping Experience can be investigated by Machine Learning such as Web Service and Web design. The target of Web design research is to determine what particular factors in a website can improve online shopping experience (Zaphiris & Kurniawan, 2007). The previous researches are mostly based on the feedback from the questionnaire, which contains limited data, but with the help of Machine Learning, we could directly collect information from customer behaviours (Zaphiris & Kurniawan, 2007). The Web Service research concentrates on the whole process of online shopping, which includes the view, purchase, delivery and quality feedback (Cardoso, Sheth, Miller, Arnold, & Kochut, 2004). That research provides a lot of complex data such as user-to-item relationship from the web viewing, user geometric information from delivery and user sentiment response from feedbacks, and those data provide the perfect resource for machine learning investigation.

More theories to fill the research gap In Chapter 5, we introduce Self Identity to distinguish online shopping customers into different groups and analyze RS based on Self Identity properties. The result shows Self Identity theory is necessary for improving RS, and the new Self Identity enhanced RS performs better than the previous one. However, we also find Self Identity theory can not cover all customer shopping behaviours, additional theory or feature need to be introduced to improve RS further. As discussed above, there are many fields in online shopping experience can be analyzed by different methods, and the associated theory is also necessary to explain and guide the research in other fields.

We conclude the limitations and implications in those four dimensions, which include both Machine Learning and Information System research areas.

It describes a lot of research trends in this field, and in our further work, we choose two directions from both theoretical field and industrial project.

To consider the longer Implications of this thesis for 5-10 years, on the one hand, I believe our proposed explainable recommender system framework will direct a new way to develop the new recommender system changing from statistics focused into theory focused. In the 5-10 years future, recommender system will be the core component of the online shopping website and highly contribute to better online shopping experience. On the other hand, the research methodology in this thesis also indicates the future of how Deep Learning technology will be included in the current Information System. The future Information System includes online shopping websites will contain more and more intelligence from Machine Learning.

6.4 Further

I address the research gap between Online Shopping Experience and Machine Learning and solve two particular problems in experiments. Since it is a big topic, there are still a lot of works to be done in this area. In further, I plan to continue on some theoretical and algorithmic works and also analyze and improve some industrial projects. Meanwhile, I will discuss the longtime future research trends in this area.

6.4.1 Image analysis with the human in the loop

In Computer Science algorithm development research field, another possible technique for improving Online Shopping Experience is to improve image analysis with the human in the loop. Most Computer Vision problems are solved by machine learning algorithms, and there is no need to build a huge image dataset to be learned by that algorithm. Rather researchers need to collect a well-labelled fashion dataset for training purpose. The quality of that dataset determined the accuracy of the computer vision model to a certain degree. However, the collection of that dataset usually is expensive and time-consuming. Especially in the fashion and clothing industry, the product and style are changing ever year, and fashion companies update their dataset frequently. To resolve this issue, a humans in the loop method is proposed (Branson et al., 2010; Mensink et al., 2011). In this method, humans' answers are collected for some specially designed questions, and these questions are formed as human knowledge to enrich the model. Compared with the previous algorithm, the Humans in the loop method use less dataset and get more intelligent results in a dynamic way.

The current progress for humans in the loop methods only have been widely used in animal datasets (Branson et al., 2010) or unfamiliar classes (Wah & Belongie, 2013). There are not any works on Online Shopping Experience mainly because the feedbacks on product items are different among different customer groups, which is not like those structured feedbacks on animals. To improve the humans in the loop methods for the Online Shopping, more feedbacks from different customers groups could be adopted in the algorithm. The past marketing research findings on customer segmentation could be considered to apply to humans in the loop methods. The integration of previous marketing theories and information systems theories is expected to contribute to the humans in the loop methods.

6.4.2 Roommate Recommendation System

Considering the industrial requirements, roommate recommendation could be a possible further trend. The online property renting is a specific field of online shopping which attracted a lot of companies to invest in recent years (Calonge, 2009). The room renting system for students is recently in high demand due to the increasing number of international students, and for properties with many rooms, those students usually need to find some others to share. However, the lack of social activity and friends for new international students becomes an important resistance for them to find suitable roommates. In this case, Road51 raised an industrial project which is to build an online roommate matching system. The company provides datasets about property features and previous renting information and wants to build a smart personalize Roommate Recommendation system. Compared with Online Shopping Recommendation system, the difficulty of this project is in three aspects.

Complex Data recourses Different from Online Recommendation System, which data only include products and customers, the Roommate Recommendation system need to consider the information from different dimensions. The first dimension is personalization which includes the personal requirement and its social preference and features. The second dimension is the property features, including price, size, and so on. The third dimension is the geometric features which indicate the geometric relationship to popular locations. The last dimension is the community information which can be collected from the City council, including the crime rate, the average age for that property located area. Those complex data need to be carefully considered and used in the roommate recommender system.

Data limitation The online shopping behaviour can happen at any time anywhere that makes Online Shopping website collected a huge amount of data for training model and predicting. However, the process of finding roommate, compared with online shopping, happen in limited times which can not provide enough data for the recommendation. Moreover, we need to apply more Information Science analyze to predict their preference.

Value for investment As an industrial project, we need to consider the trade-off between the investment and efficiency of the system. A complex recommendation system usually needs more server resources and more funds for data collection, but those additional investments may not result in the corresponding reciprocation. In this case, during the system design, we need to balance the difficulty of the system and the investment for that system.

The three difficulties differ Roommate Recommendation System to normal Online Shopping Recommendation System, but Self Identity and deep feature based recommend methods are still necessary for this project. In this case, by solving the three difficulties, we plan to take all the information into the classic recommendation model and produce real-time recommendation result.

6.4.3 Future Research Trend

The development of Deep Learning achieves a lot of sightings during the past five years, but the research about how Deep Learning can be used in Information Systems and how the customer will feedback to the new algorithms is rarely discussed. However, I believe, there will be more and more researchers working on this topic in the next five to ten years. Generally, the breakthrough of those researches would happen in two research trends.

Privacy and Security The problems of privacy and security of personal information has been discussed a lot in IS research, but rarely mentioned in Deep Learning research. Deep Learning models usually need a lot of data for training, and most of the data contains personal information which may be presented somehow in the model. For example, for a personalized recommendation system, the privacy of shopping history may be leaked by analyzing the prediction in that system. Those kinds of privacy and security control should be carefully discussed in Information System research, and I believe the problem must be solved in the future.

Creativity and Imagination The current Information System researches are mostly focused on how to use and aggregate existing information for customers efficiently. In recent research, with the help of GAN, the Deep Learning network has the potential to create new resources according to some patterns. That means the Deep Learning is not just analyzing data but also creating information. In Information System research, how that Creativity and Imagination from Deep learning can be used to help customers would be another research trends in the future.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... others (2016). Tensorflow: a system for large-scale machine learning. In *Osdi* (Vol. 16, pp. 265–283).

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 & Data Engineering*(6), 734–749.

Aghekyan-Simonian, M., Forsythe, S., Kwon, W. S., & Chattaraman, V. (2012). The role of product brand image and online store image on perceived risks and online purchase intentions for apparel. *Journal of Retailing and Consumer Services*, *19*(3), 325–331.

Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *28*(12), 2037–2041.

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

Aljukhadar, M., Senecal, S., & Ouellette, D. (2010). Can the media richness of a privacy disclosure enhance outcome? A multifaceted view of trust in rich media environments. *International Journal of Electronic Commerce*, *14*(4), 103–126.

Amazon, E. (2015). Amazon web services. *Available in: http://aws. amazon. com/es/ec2/(November 2012).*

Andersen, R., Borgs, C., Chayes, J., Feige, U., Flaxman, A., Kalai, A., ... Tennenholtz, M. (2008). Trust-based recommendation systems: an axiomatic approach. In *Proceedings of the 17th international conference on world wide web* (pp. 199–208).

Arnold, M. J., Reynolds, K. E., Ponder, N., & Lueg, J. E. (2005). Customer delight in a retail context: investigating delightful and terrible shopping experiences. *Journal of Business Research*, *58*(8), 1132–1145.

Banko, M., Cafarella, M. J., Soderland, S., Broadhead, M., & Etzioni, O. (2007). Open information extraction for the web. In *Ijcai* (Vol. 7, pp. 2670–2676).

Bauer, H. H., Falk, T., & Hammerschmidt, M. (2006). eTransQual: A transaction process-based approach for capturing service quality in online shopping. *Journal of Business Research*, *59*(7), 866–875.

Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Vaughan, J. W. (2010). A theory of learning from different domains. *Machine learning*, *79*(1-2), 151–175.

Bengio, Y. (2009). Learning deep architectures for ai. *Foundations & Trends in Machine Learning*, *2*(1), 1-127.

Bhatnagar, A., Misra, S., & Rao, H. R. (2000). On risk, convenience, and Internet shopping behavior. *Communications of the ACM*, *43*(11), 98–105.

Bossard, L., Dantone, M., Leistner, C., Wengert, C., Quack, T., & Van Gool, L. (2013). Apparel classification with style. , 321–335.

Branson, S., Wah, C., Schroff, F., Babenko, B., Welinder, P., Perona, P., & Belongie, S. (2010). Visual recognition with humans in the loop. In *Computer vision–eccv 2010* (pp. 438–451). Springer.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Brewer, M. B., & Gardner, W. (1996). Who is this" we"? levels of collective identity and self representations. *Journal of personality and social psychology*, *71*(1), 83.

Brown, M., Pope, N., & Voges, K. (2003). Buying or browsing? An exploration of shopping orientations and online purchase intention. *European Journal of Marketing*, *37*(11/12), 1666–1684.

caglar, E. (2009). Similarities and differences in physical self-concept of males and females during late adolescence and early adulthood. *Adolescence*, *44*(174).

Calonge, J. (2009, September 24). *Online system and method for property rental transactions, property management, and assessing performance of landlords and tenants.* Google Patents. (US Patent App. 12/381,956)

Cao, Z., Long, M., Wang, J., & Jordan, M. I. (2018). Partial transfer learning with selective adversarial networks. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 2724–2732).

Cardoso, J., Sheth, A., Miller, J., Arnold, J., & Kochut, K. (2004). Quality of service for workflows and web service processes. *Web Semantics: Science, Services and Agents on the World Wide Web*, *1*(3), 281–308.

Carter, M. (2015). Me, my self, and i (t): conceptualizing information technology identity and its implications. *Mis Quarterly*, *39*(4).

Chang, C.-C. (2008). Choice, perceived control, and customer satisfaction: the psychology of online service recovery. *CyberPsychology & Behavior*, *11*(3), 321–328.

Charng, H.-W., Piliavin, J. A., & Callero, P. L. (1988). Role identity and reasoned action in the prediction of repeated behavior. *Social Psychology Quarterly*, 303–317.

Chen, H., Gallagher, A., & Girod, B. (2012a). Describing clothing by semantic attributes. In *Computer vision–eccv 2012* (pp. 609–623). Springer.

Chen, H., Gallagher, A., & Girod, B. (2012b). Describing clothing by semantic attributes. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7574 LNCS*, 609–623. doi: 10.1007/978-3-642-33712-3_44

Chen, H., Xu, Z. J., Liu, Z. Q., & Zhu, S. C. (2006). Composite templates for cloth modeling and sketching. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *1*, 943–950.

Chen, X., Mottaghi, R., Liu, X., Fidler, S., Urtasun, R., & Yuille, A. (2014).

Detect what you can: Detecting and representing objects using holistic models and body parts. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 1971–1978).

Cheng, H.-T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., … others (2016). Wide & deep learning for recommender systems. In *Proceedings of the 1st workshop on deep learning for recommender systems* (pp. 7–10).

Childers, T. L., Carr, C. L., Peck, J., & Carson, S. (2002). Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of retailing*, *77*(4), 511–535.

Chu, C., Kim, S. K., Lin, Y.-A., Yu, Y., Bradski, G., Ng, A. Y., & Olukotun, K. (2007). Map-reduce for machine learning on multicore. *Advances in neural information processing systems*, *19*, 281.

Claesen, M., De Smet, F., Suykens, J. A., & De Moor, B. (2015). A robust ensemble approach to learn from positive and unlabeled data using svm base models. *Neurocomputing*, *160*, 73–84.

Cohen, S. G., & Ledford Jr, G. E. (1994). The effectiveness of self-managing teams: A quasi-experiment. *Human relations*, *47*(1), 13–43.

Cui, Y., Song, Y., Sun, C., Howard, A., & Belongie, S. (2018). Large scale fine-grained categorization and domain-specific transfer learning. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 4109–4118).

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Computer vision and pattern recognition, 2005. cvpr 2005. ieee computer society conference on* (Vol. 1, pp. 886–893).

de Amorim, R. C. (2012). Constrained clustering with minkowski weighted k-means. In *Computational intelligence and informatics (cinti), 2012 ieee 13th international symposium on* (pp. 13–17).

Deng, J. (2012). Large scale visual recognition. *Dissertations & Theses - Gradworks*.

Deng, J., Dong, W., Socher, R., & Li, L. J. (2009). Imagenet: A large-scale hierarchical image database. In *Computer vision and pattern recognition, 2009. cvpr 2009. ieee conference on* (p. 248-255).

Der, M., & Saul, L. K. (2012). Latent coincidence analysis: A hidden variable model for distance metric learning. In *Advances in neural information processing systems* (pp. 3230–3238).

Di, W., Wah, C., Bhardwaj, A., Piramuthu, R., & Sundaresan, N. (2013a). Style finder: Fine-grained clothing style detection and retrieval. In *Computer vision and pattern recognition workshops (cvprw), 2013 ieee conference on* (pp. 8–13).

Di, W., Wah, C., Bhardwaj, A., Piramuthu, R., & Sundaresan, N. (2013b). Style finder: Fine-grained clothing style detection and retrieval. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 8–13. doi: 10.1109/CVPRW.2013.6

Dix, A. (2009). *Human-computer interaction*. Springer.

Doolin, B., Dillons, S., Thompson, F., & Corner, J. L. (2007). Perceived risk, the Internet shopping experience and online purchasing behavior: A New Zealand perspective. *Electronic commerce: Concepts, methodologies, tools, and applications*, 324–345.

Farhadi, A., Endres, I., Hoiem, D., & Forsyth, D. (2009). Describing objects by their attributes. In *Computer vision and pattern recognition, 2009. cvpr 2009. ieee conference on* (pp. 1778–1785).

Fatma, S. (2014). Antecedents and consequences of customer experience management-a literature review and research agenda. *International Journal of Business and Commerce*, *3*(6), 32–49.

Fei-Fei, L., Fergus, R., & Perona, P. (2006). One-shot learning of object categories. *IEEE transactions on pattern analysis and machine intelligence*, *28*(4), 594–611.

Fei-Fei, L., & Perona, P. (2005). A bayesian hierarchical model for learning natural scene categories. In *Computer vision and pattern recognition, 2005. cvpr 2005. ieee computer society conference on* (Vol. 2, pp. 524–531).

Ferrari, V., & Zisserman, A. (2008). Learning visual attributes.

Fiore, A. M., Kim, J., & Lee, H.-H. (2005). Effect of image interactivity technology on consumer responses toward the online retailer. *Journal of Interactive Marketing*, *19*(3), 38–53.

Fogel, S., Averbuch-Elor, H., Cohen-Or, D., & Goldberger, J. (2019). Clustering-driven deep embedding with pairwise constraints. *IEEE computer graphics and applications*, *39*(4), 16–27.

Freiberg, J., Tracy, B., Sanborn, M., & McKain, S. (2009). Praise for The Starbucks Experience.

Frow, P., & Payne, A. (2007). Towards the 'perfect'customer experience. *Journal of Brand Management*, *15*(2), 89–101.

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine learning*, *63*(1), 3–42.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., … Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672–2680).

Granberg, D., & Holmberg, S. (1990). The intention-behavior relationship among us and swedish voters. *Social Psychology Quarterly*, 44–54.

Grewal, D., Levy, M., & Kumar, V. (2009). Customer experience management in retailing: An organizing framework. *Journal of retailing*, *85*(1), 1–14.

Guan, Y., Wei, Q., & Chen, G. (2019). Deep learning based personalized recommendation with multi-view information integration. *Decision Support Systems*, *118*, 58–69.

Ha, S., & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of Business Research*, *62*(5), 565–571.

Hashim, A., Ghani, E. K., & Said, J. (2009). Does Consumers' Demographic Profile Influence Online Shopping?: An Examination Using Fishbein's Theory/Est-Ce Que Le Profile Démographique Des Consommateurs Influence Les Achats En Ligne?: Une Expertise En Utilisant La Théorie de Fishbein. *Canadian Social Science*, *5*(6), 19.

Häubl, G., & Trifts, V. (2000). Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing science*, *19*(1), 4–21.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 770–778).

He, R., & McAuley, J. (2016). Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *proceedings of the 25th international conference on world wide web* (pp. 507–517).

Hernández, B., Jiménez, J., & José Martín, M. (2011). Age, gender and income: do they really moderate online shopping behaviour? *Online Information Review*, *35*(1), 113–133.

Hernández, B., Jiménez, J., & Martín, M. J. (2011). Age, gender and income: do they really moderate online shopping behaviour? *Online Information Review*, *35*, 113–133. doi: 10.1108/14684521111113614

Hidasi, B., Karatzoglou, A., Baltrunas, L., & Tikk, D. (2015). Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939*.

Hinton, G. E., Osindero, S., & Teh, Y. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, *18*(7), 1527-1554.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, *313*(5786), 504–507.

Ho, T. K. (1998). The random subspace method for constructing decision forests. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *20*(8), 832–844.

Hoffman, D. L., & Novak, T. P. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *The Journal of Marketing*, 50–68.

Hogg, M. A. (2001). A social identity theory of leadership. *Personality and social psychology review*, *5*(3), 184–200.

Hogg, M. A., & Terry, D. I. (2000). Social identity and self-categorization processes in organizational contexts. *Academy of management review*, *25*(1), 121–140.

Horowitz, M. J. (2012). Self-identity theory and research methods. *Journal of Research Practice*, *8*(2), 14.

Houde, S., & Hill, C. (1997). What do prototypes prototype? In *Handbook of human-computer interaction (second edition)* (pp. 367–381). Elsevier.

Hu, Y., Wang, J., Yu, N., & Hua, X.-S. (2008). Maximum margin clustering with pairwise constraints. In *Data mining, 2008. icdm'08. eighth ieee international conference on* (pp. 253–262).

Hu, Y., Yi, X., & Davis, L. S. (2015a). Collaborative fashion recommendation: A functional tensor factorization approach. In *Proceedings of the 23rd annual acm conference on multimedia conference* (pp. 129–138).

Hu, Y., Yi, X., & Davis, L. S. (2015b). Collaborative fashion recommendation: A functional tensor factorization approach. In *Proceedings of the 23rd acm international conference on multimedia* (pp. 129–138). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2733373.2806239` doi: 10.1145/2733373.2806239

Iwata, T., Wanatabe, S., & Sawada, H. (2011). Fashion coordinates recommender system using photographs from fashion magazines. In *Ijcai proceedings-international joint conference on artificial intelligence* (Vol. 22, p. 2262).

Jackson, P., & Moulinier, I. (2007). *Natural language processing for online applications: Text retrieval, extraction and categorization* (Vol. 5). John Benjamins Publishing.

Jacobs, P. S. (2014). *Text-based intelligent systems: Current research and practice in information extraction and retrieval.* Psychology Press.

Jagadeesh, V., Piramuthu, R., Bhardwaj, A., Di, W., & Sundaresan, N. (2014). Large scale visual recommendations from street fashion images. In *Proceedings of the 20th acm sigkdd international conference on knowledge* *discovery and data mining* (pp. 1925–1934).

Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). *Recommender systems: an introduction*. Cambridge University Press.

Jerry, B. (2005). *Discrete event system simulation*. Pearson Education India.

Johnson, R. E., Venus, M., Lanaj, K., Mao, C., & Chang, C.-H. (2012). Leader identity as an antecedent of the frequency and consistency of transformational, consideration, and abusive leadership behaviors. *Journal of Applied Psychology*, *97*(6), 1262.

Jones, M. A. (1999). Entertaining shopping experiences: an exploratory investigation. *Journal of retailing and consumer services*, *6*(3), 129–139.

Kalantidis, Y., Kennedy, L., & Li, L.-J. (2013). Getting the look: clothing recognition and segmentation for automatic product suggestions in everyday photos. , 105–112.

Kaynama, S. A., & Black, C. I. (2000). A proposal to assess the service quality of online travel agencies: An exploratory study. *Journal of professional services marketing*, *21*(1), 63–88.

Ke, Y., & Sukthankar, R. (2004). Pca-sift: A more distinctive representation for local image descriptors. In *Computer vision and pattern recognition, 2004. cvpr 2004. proceedings of the 2004 ieee computer society conference on* (Vol. 2, pp. II–506).

Kimery, K. M., & McCord, M. (2002). Third-party assurances: Mapping the road to trust in e-retailing. *JITTA: Journal of Information Technology Theory and Application*, *4*(2), 63.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kouloumpis, E., Wilson, T., & Moore, J. D. (2011). Twitter sentiment analysis: The good the bad and the omg! *Icwsm*, *11*(538-541), 164.

Kovashka, A., & Grauman, K. (2013). Attribute adaptation for personalized image search. In *Computer vision (iccv), 2013 ieee international conference on* (pp. 3432–3439).

Kranzbühler, A.-M., Kleijnen, M. H., Morgan, R. E., & Teerling, M. (2018). The multilevel nature of customer experience research: an integrative review and research agenda. *International Journal of Management Reviews*, *20*(2), 433–456.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *International conference on neural information processing systems* (p. 1097-1105).

Kumar, V., Khattar, D., Gupta, S., Gupta, M., & Varma, V. (2017). Deep neural architecture for news recommendation. In *Working notes of the 8th international conference of the clef initiative, dublin, ireland. ceur workshop proceedings.*

Kuo, H.-M., & Chen, C.-W. (2011). Application of quality function deployment to improve the quality of internet shopping website interface design. *International Journal of Innovative Computing, Information and Control*, *7*(1), 253–268.

Kuo, H.-M., Chen, C.-W., & Hsu, C.-H. (2012). A study of a b2c supporting interface design system for the elderly. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *22*(6), 528–540.

Lai, C.-Y., Shih, D.-H., Chiang, H.-S., & Chen, C.-C. (2009). The key factors of influence consumer online shopping behavior: using the iqa approach. In *Proceedings of the 8th wseas international conference on e-activities and information security and privacy* (pp. 286–291).

Lampert, C. H., Nickisch, H., & Harmeling, S. (2009). Learning to detect unseen object classes by between-class attribute transfer. In *Computer vision and pattern recognition, 2009. cvpr 2009. ieee conference on* (pp. 951–958).

Law, A. M., Kelton, W. D., & Kelton, W. D. (1991). *Simulation modeling and analysis* (Vol. 2). McGraw-Hill New York.

Lazebnik, S., Schmid, C., & Ponce, J. (2006). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Computer vision and pattern recognition, 2006 ieee computer society* *conference on* (Vol. 2, pp. 2169–2178).

Le, Q. V. (2013). Building high-level features using large scale unsupervised learning. In *Ieee international conference on acoustics, speech and signal processing* (p. 8595-8598).

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

LeCun, Y., Jackel, L., Bottou, L., Brunot, A., Cortes, C., Denker, J., … others (1995). Comparison of learning algorithms for handwritten digit recognition. In *International conference on artificial neural networks* (Vol. 60, pp. 53–60).

Lee, G.-G., & Lin, H.-F. (2005). Customer perceptions of e-service quality in online shopping. *International Journal of Retail & Distribution Management*, *33*(2), 161–176.

Lee, S. A., & Jeong, M. (2014). Enhancing online brand experiences: An application of congruity theory. *International Journal of Hospitality Management*, *40*, 49–58.

Leventhal, R. C., Mascarenhas, O. A., Kesavan, R., & Bernacchi, M. (2006). Lasting customer loyalty: a total customer experience approach. *Journal of Consumer Marketing*, *23*(7), 397–405.

Levin, A. M., Levin, I. P., & Weller, J. A. (2005). A multi-attribute analysis of preferences for online and offline shopping: Differences across products, consumers, and shopping stages. *Journal of Electronic Commerce Research*, *6*(4), 281.

Li, H., Aham-Anyanwu, N., Tevrizci, C., & Luo, X. (2015). The interplay between value and service quality experience: e-loyalty development process through the eTailQ scale and value perception. *Electronic Commerce Research*, *15*(4), 585–615.

Li, Y., Liu, T., Jiang, J., & Zhang, L. (2016). Hashtag recommendation with topical attention-based lstm..

Liao, Z., & Cheung, M. T. (2001). Internet-based e-shopping and consumer attitudes: an empirical study. *Information & Management*, *38*(5), 299–

306.

Linden, G., Smith, B., & York, J. (2003). Amazon. com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE*, *7*(1), 76–80.

Liu, Q., Wu, S., & Wang, L. (2017). Deepstyle: Learning user preferences for visual recommendation. In *Proceedings of the 40th international acm sigir conference on research and development in information retrieval* (pp. 841–844).

Liu, S., Feng, J., Song, Z., Zhang, T., Lu, H., Xu, C., & Yan, S. (2012). "Hi , Magic Closet , Tell Me What to Wear !". *ACM Multimedia*, 619–628. doi: 10.1145/2393347.2393433

Liu, X., Song, M., Tao, D., Liu, Z., Zhang, L., Chen, C., & Bu, J. (2013). Semi-supervised node splitting for random forest construction. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 492–499).

Liu, Z., Luo, P., Qiu, S., Wang, X., & Tang, X. (2016). Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 1096–1104).

Loiacono, E. T., Watson, R. T., & Goodhue, D. L. (2002). WebQual: A measure of website quality. *Marketing theory and applications*, *13*(3), 432–438.

Loni, B., Cheung, L. Y., Riegler, M., Bozzon, A., Gottlieb, L., & Larson, M. (2014). Fashion 10000: an enriched social image dataset for fashion and clothing. In *Proceedings of the 5th acm multimedia systems conference* (pp. 41–46).

Lops, P., De Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In *Recommender systems handbook* (pp. 73–105). Springer.

Luo, J., Ba, S., & Zhang, H. (2012). The effectiveness of online shopping characteristics and well-designed websites on satisfaction. *Mis Quar-*

terly, *36*(4), 1131–1144.

Machine learning /. (1997). McGraw-Hill,.

Manning, & ChristopherD. (2008). *Introduction to information retrieval*. MIT Press,.

Manning, C. D. (1999). *Foundations of statistical natural language processing*. MIT Press,.

Maxwell, J. A. (2008). Designing a qualitative study. *The SAGE handbook of applied social research methods*, *2*, 214–253.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: a trust building model. *The Journal of Strategic Information Systems*, *11*(3), 297–323.

Mensink, T., Verbeek, J., & Csurka, G. (2011). Learning structured prediction models for interactive image labeling. In *Computer vision and pattern recognition (cvpr), 2011 ieee conference on* (pp. 833–840).

Meyer, C., & Schwager, A. (2007). Customer Experience. *Harvard business review*, 1–11.

Min, S., Lee, B., & Yoon, S. (2017). Deep learning in bioinformatics. *Briefings in bioinformatics*, *18*(5), 851–869.

Minsky, M., & Papert, S. (1969). Perceptrons. *American Journal of Psychology*, *84*(3), 449–452.

Miyazaki, A. D., & Fernandez, A. (2001). Consumer perceptions of privacy and security risks for online shopping. *Journal of Consumer affairs*, *35*(1), 27–44.

Mnih, V., & Hinton, G. E. (2012). Learning to label aerial images from noisy data. In *Proceedings of the 29th international conference on machine learning (icml-12)* (pp. 567–574).

Moorman, C., Zaltman, G., & Deshpande, R. (1992). Relationships between providers and users of market research: The dynamics of trust. *Journal of marketing research*, *29*(3), 314–328.

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *The journal of marketing*, 20–38.

Mudaly, V., & De Villiers, M. (2004). Mathematical modeling and proof. In *Proceedings of the 10th conference of association for mathematics education of south africa, university of the nort-west, potchefstroom.*

Nguyen, N., & Caruana, R. (2008). Improving classification with pairwise constraints: a margin-based approach. In *Machine learning and knowledge discovery in databases* (pp. 113–124). Springer.

Nicolas, C. (2004, December 16). *Mobile phone sample survey method.* Google Patents. (US Patent App. 10/460,374)

Noroozi, M., Vinjimoor, A., Favaro, P., & Pirsiavash, H. (2018). Boosting self-supervised learning via knowledge transfer. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 9359–9367).

Novak, T. P., Hoffman, D. L., & Yung, Y.-F. (2000). Measuring the customer experience in online environments: A structural modeling approach. *Marketing science*, *19*(1), 22–42.

Nurra, C., & Oyserman, D. (2018). From future self to current action: An identity-based motivation perspective. *Self and identity*, *17*(3), 343–364.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, *42*(3), 145–175.

Olshen, L., Stone, C. J., et al. (1984). Classification and regression trees. *Wadsworth International Group*, *93*(99), 101.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, *22*(10), 1345–1359.

Papacharissi, Z. (2010). *A networked self: Identity, community, and culture on social network sites*. Routledge.

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). ES-QUAL a multiple-item scale for assessing electronic service quality. *Journal of service research*, *7*(3), 213–233.

Park, C.-H., & Kim, Y.-G. (2003). Identifying key factors affecting consumer purchase behavior in an online shopping context. *International Journal of Retail & Distribution Management*, *31*(1), 16–29.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Perea y Monsuwé, T., Dellaert, B. G., & De Ruyter, K. (2004). What drives consumers to shop online? a literature review. *International journal of service industry management*, *15*(1), 102–121.

Peyre, J., Sivic, J., Laptev, I., & Schmid, C. (2017). Weakly-supervised learning of visual relations. In *Proceedings of the ieee international conference on computer vision* (pp. 5179–5188).

Poon, A., Maltzman, R., & Taylor, J. (2012, September 25). *Method and system to recommend further items to a user of a network-based transaction facility upon unsuccessful transacting with respect to an item.* Google Patents. (US Patent 8,275,673)

Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, *2016*(1), 67.

Quadrana, M., Cremonesi, P., & Jannach, D. (2018). Sequence-aware recommender systems. *ACM Computing Surveys (CSUR)*, *51*(4), 66.

Quan, S., & Wang, N. (2004). Towards a structural model of the tourist experience: An illustration from food experiences in tourism. *Tourism management*, *25*(3), 297–305.

Ranganathan, C., & Ganapathy, S. (2002). Key dimensions of business-to-consumer web sites. *Information & Management*, *39*(6), 457–465.

Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: introduction and challenges. In *Recommender systems handbook* (pp. 1–34). Springer.

Rijmenam, M. (2016). *How amazon is leveraging big data.* Retrieved 2016-02-21, from `https://datafloq.com/read/amazon-leveraging-big-data/517`

Robillard, M. P., Walker, R. J., & Zimmermann, T. (2010). Recommendation systems for software engineering. *Software, IEEE*, *27*(4), 80–86.

Rose, S., Hair, N., & Clark, M. (2011). Online customer experience: A review of the business-to-consumer online purchase context. *International Journal of Management Reviews*, *13*(1), 24–39.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of management review*, *23*(3), 393–404.

Sarkar, A. (2011). Impact of utilitarian and hedonic shopping values on individual's perceived benefits and risks in online shopping. *International management review*, *7*(1), 58–65.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on world wide web* (pp. 285–295).

Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. *Data mining and knowledge discovery*, *5*(1-2), 115–153.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, *61*, 85–117.

Schmitz, M., Bart, R., Soderland, S., Etzioni, O., et al. (2012). Open language learning for information extraction. In *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning* (pp. 523–534).

Schoder, D., & Yin, P.-L. (2000). Building firm trust online. *Communications of the ACM*, *43*(12), 73–79.

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 815–823).

Seo, S., Huang, J., Yang, H., & Liu, Y. (2017). Representation learning of users and items for review rating prediction using attention-based convolutional neural network. In *3rd international workshop on machine learning methods for recommender systems (mlrec)(sdm'17).*

Shang, F., Jiao, L., & Liu, Y. (2012). Integrating spectral kernel learning and constraints in semi-supervised classification. *Neural processing letters*, *36*(2), 101–115.

Shi, Y., Otto, C., & Jain, A. K. (2018). Face clustering: representation and pairwise constraints. *IEEE Transactions on Information Forensics and Security*, *13*(7), 1626–1640.

Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., … Moore, R. (2013). *Real-time human pose recognition in parts from single depth images*. Springer Berlin Heidelberg.

Sigurbjörnsson, B., & Van Zwol, R. (2008). Flickr tag recommendation based on collective knowledge. In *Proceedings of the 17th international conference on world wide web* (pp. 327–336).

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Socher, R., Ganjoo, M., Manning, C. D., & Ng, A. (2013). Zero-shot learning through cross-modal transfer. In *Advances in neural information processing systems* (pp. 935–943).

Sohn, C., & Tadisina, S. K. (2008). Development of e-service quality measure for internet-based financial institutions. *Total Quality Management*, *19*(9), 903–918.

Sparks, P. (2000). Subjective expected utility-based attitude–behavior models: The utility of self-identity.

Stets, J. E., & Burke, P. J. (2003). A sociological approach to self and identity. *Handbook of self and identity*, *128152*.

Stryker, S., & Burke, P. J. (2000). The Past, Present, and Future of an Identity Theory. *Social Psychology Quarterly*. doi: 10.2307/2695840

Suki, N. M., & Suki, N. M. (2007). Online buying innovativeness: effects of perceived value, perceived risk and perceived enjoyment. *International journal of Business and Society*, *8*(2), 81–93.

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first aaai conference on artificial intelligence.*

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., … Rabinovich, A. (2015). Going deeper with convolutions. In *Computer vision and pattern recognition* (p. 1-9).

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. In *International conference on artificial neural networks* (pp. 270–279).

Tan, M., & Le, Q. V. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*.

Terry, D. J., Hogg, M. A., & White, K. M. (1999). The theory of planned behaviour: self-identity, social identity and group norms. *British Journal of Social Psychology*, *38*(3), 225–244.

Thoits, P. A., & Virshup, L. K. (1997). Me's and we's. *Self and identity: Fundamental issues*, 106–133.

Thusoo, A., Sarma, J. S., Jain, N., Shao, Z., Chakka, P., Anthony, S., … Murthy, R. (2009). Hive: a warehousing solution over a map-reduce framework. *Proceedings of the VLDB Endowment*, *2*(2), 1626–1629.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: computation in neural systems*, *14*(3), 391–412.

Trautwein, U., Lüdtke, O., Marsh, H. W., & Nagy, G. (2009). Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of Educational Psychology*, *101*(4), 853.

Tu, Y., Lin, Y., Wang, J., & Kim, J.-U. (2018). Semi-supervised learning with generative adversarial networks on digital signal modulation classification. *Comput. Mater. Continua*, *55*(2), 243–254.

van Baalen, M. (2016). Deep matrix factorization for recommendation.

Van der Werff, E., Steg, L., & Keizer, K. (2013). The value of environmental self-identity: The relationship between biospheric values, environmental self-identity and environmental preferences, intentions and behaviour. *Journal of Environmental Psychology*, *34*, 55–63.

Van Meteren, R., & Van Someren, M. (2000). Using content-based filtering for recommendation. In *Proceedings of the machine learning in the new information age: Mlnet/ecml2000 workshop* (pp. 47–56).

Varma, M., & Zisserman, A. (2005). A statistical approach to texture classification from single images. *International Journal of Computer Vision*, *62*(1-2), 61–81.

Vedaldi, A., & Fulkerson, B. (2008). *VLFeat: An open and portable library of computer vision algorithms.* http://www.vlfeat.org/.

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of retailing*, *85*(1), 31–41.

Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P. A. (2008). Extracting and composing robust features with denoising autoencoders. In *International conference on machine learning* (p. 1096-1103).

Wagstaff, K., Cardie, C., Rogers, S., Schrödl, S., et al. (2001). Constrained k-means clustering with background knowledge. In *Icml* (Vol. 1, pp. 577–584).

Wah, C., & Belongie, S. (2013). Attribute-based detection of unfamiliar classes with humans in the loop. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 779–786).

Wang, S., Wang, Y., Tang, J., Shu, K., Ranganath, S., & Liu, H. (2017). What your images reveal: Exploiting visual contents for point-of-interest recommendation. In *Proceedings of the 26th international conference on world wide web* (pp. 391–400).

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, *3*(1), 9.

Whitmarsh, L., & O'Neill, S. (2010). Green identity, green living? the role of pro-environmental self-identity in determining consistency across diverse pro-environmental behaviours. *Journal of Environmental Psychology*, *30*(3), 305–314.

Witten, I. H., & Frank, E. (2005). *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Wolfinbarger, M., & Gilly, M. (2002). . comQ: dimensionalizing, measuring, and predicting quality of the e-tail experience. *Marketing Science Institute Report*(02-100).

Wolfinbarger, M., & Gilly, M. C. (2003). etailq: dimensionalizing, measuring and predicting etail quality. *Journal of retailing*, *79*(3), 183–198.

Xiang, S., Nie, F., & Zhang, C. (2008). Learning a mahalanobis distance metric for data clustering and classification. *Pattern Recognition*, *41*(12), 3600–3612.

Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated residual transformations for deep neural networks. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 1492–1500).

Xu, J. D., Benbasat, I., & Cenfetelli, R. T. (2013). Integrating service quality with system and information quality: an empirical test in the e-service context. *Mis Quarterly*, *37*(3), 777–794.

Yamaguchi, K., Kiapour, M. H., Ortiz, L. E., & Berg, T. L. (2012). Parsing clothing in fashion photographs. In *Computer vision and pattern recognition (cvpr), 2012 ieee conference on* (pp. 3570–3577).

Yan, R., Zhang, J., Yang, J., & Hauptmann, A. G. (2006). A discriminative learning framework with pairwise constraints for video object classification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *28*(4), 578–593.

Yang, Y., & Ramanan, D. (2013). Articulated human detection with flexible mixtures of parts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *35*(12), 2878–2890.

Ying, R., He, R., Chen, K., Eksombatchai, P., Hamilton, W. L., & Leskovec, J. (2018). Graph convolutional neural networks for web-scale recommender systems. In *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining* (pp. 974–983).

Yue, Y., Wang, C., El-Arini, K., & Guestrin, C. (2014). Personalized collaborative clustering. In *Proceedings of the 23rd international conference on world wide web* (pp. 75–84).

Zaphiris, P., & Kurniawan, S. (2007). *Human computer interaction research in web design and evaluation*. IGI Global.

Zeng, H., & Cheung, Y.-m. (2012). Semi-supervised maximum margin clustering with pairwise constraints. *Knowledge and Data Engineering, IEEE Transactions on*, *24*(5), 926–939.

Zhang, S., Yao, L., & Sun, A. (2017). Deep learning based recommender system: A survey and new perspectives. *arXiv preprint arXiv:1707.07435*.

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), 5.

Zhang, Y., Ai, Q., Chen, X., & Croft, W. B. (2017). Joint representation learning for top-n recommendation with heterogeneous information sources. In *Proceedings of the 2017 acm on conference on information and knowledge management* (pp. 1449–1458).

Zhu, S., Urtasun, R., Fidler, S., Lin, D., & Change Loy, C. (2017). Be your own prada: Fashion synthesis with structural coherence. In *Proceedings of the ieee international conference on computer vision* (pp. 1680–1688).

Zhu, X., Loy, C. C., & Gong, S. (2015). Constrained clustering with imperfect oracles.

Zhukov, D., Alayrac, J.-B., Cinbis, R. G., Fouhey, D., Laptev, I., & Sivic, J. (2019). Cross-task weakly supervised learning from instructional videos. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 3537–3545).

Zviran, M., Glezer, C., & Avni, I. (2006). User satisfaction from commercial web sites: The effect of design and use. *Information & Management*, *43*(2), 157–178.