A decision support system (DSS) is used for analysing a situation and making decisions. The goal of this research is to mine a large set of heterogeneous audiology data and create a DSS to help audiologists to choose between an ITE or BTE hearing aid. Although in many cases such a choice is clear cut, but at other times this system could be used as a second opinion to provide the hearing aid. A number of data mining techniques, such as clustering of audiograms, association analysis of variables (such as, age, gender, diagnosis, mask, mould and free text keywords) using contingency tables and principal component analysis or an algorithm to merge the data were used to find candidate variables to be combined into a DSS. The DSS was trained using the techniques of logistic regression, Naïve Bayesian analysis and Bayes network, and these systems were tested and validated on test data to which the techniques produced the better results. This DSS takes air and bone conduction frequencies, age, gender, diagnosis, mask, and some free text words associated with a patient as input and gives as the output a decision as to whether the patient would be more likely to prefer an ITE or a BTE hearing aid type.

The highest agreement between predicted results and actual hearing aid type in the data were obtained using Bayesian network, with 91 to 94 percent similarity overall, with a precision of 0.9 and 0.96 for ITE and BTE. The reason for this might be that the Bayesian network also considers interaction between variables while the other two techniques (logistic regression and Naïve Bayesian analysis) consider only the individual variables. One of the important features of this DSS is that once the final choice of hearing aid type is predicted, the decision of whether to fit a BTE or an ITE aid or a tinnitus masker.

The theoretical upper bound of classifier performance is the inter-annotator agreement (Altmann, 1991), in this case the rate at which between 2 expert audiologists would assign the same hearing aid to the same patient. Unfortunately, this type of data was not included in the audiology database.

Why data mining?
- Tremendous growth in the amount of data produced in the medical domain (Aniladhu, 2009) and each year the amount of data is increased and new terms emerge.
- The amount of audiology data used in this research is much more than the most of the studies been done in the past.
- Data mining of audiology data has certain limitations with traditional audiology research, such as small data samples, data related variables while the other two techniques (logistic regression and Naïve Bayesian analysis) consider only the individual variables.
- The availability of a large NHS audiology facility data (which is used in this research), oversteps the need for surveys / interviews / new hearing tests.

Audiology data repository
- A large NHS audiology facility
- 180,000 records, 23,000 patients
- Heterogeneous records
- e.g., IOS/ISO-95/95/95/95 (60/60/70/70)
- Contingency data e.g., [Ml,MITTINTUS, NHS, BTE]18
- Free text notes e.g., [IMP5, TAKEN FOR BINAURAL AIDS]

Hypothesis:
The data mining of heterogeneous audiology data will give the opportunity to discover new features and associations that will be helpful for audiologists in their work, in particular by developing a decision support system (DSS) to classify patients according to the following research questions:

- Research Questions:
  - Which factors influence the choice of ITE (in the ear) as opposed to BTE (behind the ear) hearing aid?
  - For patients diagnosed with tinnitus, which factors influence the decision whether to fit a tinnitus masker?

An overall work diagram of data mining of audiology data

Motivation for the research
- People with hearing loss are found to have:
  - Serious depression due to loss of hearing
  - Problems in speech and communication
- Avoiding public places due to unsafe environment like heavy traffic.
- In a study it was found that (Bertoli et al., 2009)
  - 85% people using their hearing aids
  - Regularly
  - 12% occasionally and
  - 5% never, as they may or may not be happy with their hearing aids.
- About 15% of the world’s population experiences the symptoms of tinnitus, and this increase is found, frequently and
  - Approximately 5% of the adult population in the UK experience persistent or troublesome tinnitus.
- There is the need for a second opinion in difficult-to-decide cases for hearing aid type and the use of a tinnitus masker with an explanation how that decision was arrived at.

The factors influencing the choice of ITE/BTE aids and tinnitus masker found in the literature are audiograms, age and gender. These factors were also found experimentally to affect the decision of whether to fit a BTE or an ITE aid or a tinnitus masker.

The Bayesian network considers the interaction between variables, while logistic regression and Naïve Bayesian analysis consider only the individual variables in isolation.

Among all three techniques, Bayesian network produced the better results for ITE hearing aids, while for tinnitus maskers, there was not enough data.

For the analysis of results obtained from DSS precision, recall and F-measures were used.

The final model was tested using 5-fold cross validation for logistic regression, Naïve Bayesian analysis and the Bayesian network, to choose the best algorithm to choose whether to fit an ITE or a BTE hearing aid with the results shown in Table 1 and Table 2.

Table 1: Overall similarity of results with logistic regression, Naïve Bayesian analysis and Bayesian network.

<table>
<thead>
<tr>
<th>Overall similarity</th>
<th>Percentage range</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITE</td>
<td>BTE</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>0.79</td>
</tr>
<tr>
<td>ZeroR</td>
<td>0.67</td>
</tr>
<tr>
<td>Naïve</td>
<td>0.85</td>
</tr>
<tr>
<td>Bayesian analysis</td>
<td>0.78</td>
</tr>
<tr>
<td>Bayesian network</td>
<td>0.91</td>
</tr>
</tbody>
</table>

References

Decision support system to help choose between an ITE or a BTE hearing aid
Muhammad Naveed Anwar, Michael Philip Oakes
Department of Computing, Engineering & Technology, University of Sunderland, Sunderland, UK
(mna2606@yahoo.com, Michael.Oakes@sunderland.ac.uk)