CHAPTER 8

CONCLUSION

This chapter presents the achievement of result through the review of literature and the recommendation for the future research.

Information technology in medical field had spread its target to support physicians, health related management and Classificationist to handle the whole process of medicine with various types of data. The proposed decision tree method assists in achieving solutions to satisfy multiple criteria in the health care services. There is a long way towards an effective classification method which handles various structures of data. But the proposed methods can be regarded as an important step towards the objectives.

The goal of this work is to evaluate approaches to supervised training using a combination of different classification methods with C4.5 decision tree as the base algorithm. Extensive experiment have been conducted for state of art feature selectors that lead to several important conclusions which justifies the problem of crispness of data, randomness of data and inclusion of new subset features over the old.

C4.5 and decision tree algorithm is applicable for this problem and should be more widely used in this domain. The main benefit of C4.5 is the ability to handle multidimensional data automatically. Using C4.5
implementation decision tree is constructed based on emerging patterns which is a means of discovering distinguish collection of data set, a powerful method for constructing accurate classifiers.

From the parameter comparison among four datasets it is concluded Weighted Decision Tree with Emerging Pattern and Feature Ranking algorithm (WDT-EPFR) has better performance (accuracy) than Weighted Decision tree algorithm. Feature selection technique is used for selecting subset of relevant features from the data set to build robust learning models. Classification accuracy is improved by removing most irrelevant and redundant features.

Fuzzy rough set framework (RFRS) is developed using GFRS rule-based classifier and C4.5 decision tree. The framework is based on the perception of random fuzzy-rough set. Thus it solves the problem about data perturbation due to randomness. The developed robust model builds the classifier to keep the discernibility information in the original data sets invariant. The fuzzy lower approximation operator is used in attribute value reduction.

In the final work the irrelevant features or subset features are selected using greedy hill climbing, then the Neuro Network (NN) learning algorithm is proposed, then trained neural network ensemble is employed to generate a new training set through replacing the desired class labels of the original training examples, with those output from the trained NN ensemble.
FUTURE WORK

This research focuses on classification using Decision Trees with three different approaches. Some future enhancements would be

- Propose an ensemble of flexible neural trees. This flexible neural network allows the input variables selection, over-layer connections and different activation functions for different nodes.
- Complementary ways to handle continuous attributes.
- Various machine learning algorithms can be encapsulated for better learning process.
- In the presence of multimodality decision tree can be leading to high error rates. The partitioning of the data using clustering before classification can lead to simpler trees and lower error rates.
- The unsupervised feature selection approaches can be used for the datasets which contains more unlabeled data. Here feature interaction and certainty of features are used for selecting the features.
- Rule pruning strategies can be integrated with our decision tree classifier. Ontology based taxonomy is considered while forming decision tree, this is very helpful in sub classification of decision variables.
- A fuzzy membership values of fuzzy rough set can be automatically tuned by optimization techniques, it will improve the feature selection approach.