CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 CONCLUSION

In this research, data mining approach was investigated to identify NFR from FR Documents. Initial investigations were conducted using Meta learners to identify NFR from FR Documents.

In the proposed method, the Promise database repository was used to investigate Bagging and Logitboost classification algorithms. 57 words based on their importance were extracted from the requirement document for the Data Mining operation. Precision of up to 83.1% were obtained which may not be sufficient for a fully automated system.

In the next level of investigations, it was proposed to use Soft Computing techniques to improve the Precision and Classification Accuracy. Multi Layer Perceptron (MLP) and Generalized Feed Forward Neural Networks (GFFNN) were used to investigate the efficiency to classify NFR from FR. Experimental results showed that Genetic Feed Forward Neural Networks with Sigmoid Activation achieved the Best Accuracy of 85.26%. The Classification Accuracy obtained was in the range of 73.72% to 85.26% showing an improvement of 3.5% over Meta Learning algorithms that were tested. Further work needs to be done so that the Precision of all the classes of
NFR are high. Thus, optimization techniques were incorporated in the final stages of research to further improve the Classification Accuracy.

A Neural Network classifier was proposed and implemented for retrieval of Non Functional Requirements (NFR) from Functional Requirements (FR). It was proposed to implement a Gaussian Radial Basis Function in the Hidden Layer and optimize the same using Genetic Algorithm (GA). The proposed method was compared with existing Information Retrieval (IR) techniques including Logitboost, Bagging and Multi Layer Perceptron (MLP) Neural Network. Classification accuracy obtained was in the range of 81.25% to 90.23% showing an improvement of 9.54%.

Features were extracted from the requirement documents using the proposed feature extraction system. The proposed feature extraction builds a NFR repository is built and features are selected based on the repository to retrieve NFR from documents for classification. To further improve the classification accuracy, a Hybrid Genetic Algorithm Continuous Ant Colony Optimization (HGACACO) which incorporates both Genetic Algorithm and Ant Colony Optimization (ACO) is proposed. The proposed techniques are benchmarked with the dataset from promise repository and evaluated using a new dataset. HGACACO GFFNN classifier with proposed feature extraction technique achieves a classification accuracy of 93.59% for the promise repository dataset and an accuracy of 91.06% for the new dataset.
6.2 FUTURE WORK

Further work needs to be done by using Neural Network based retrieval system and preprocessing the data with Singular Value Decomposition. This work can be further extended by testing the proposed method with

- IET requirement document used in J. Cleland-Huang et al 2007

- Study the performance of Radial Basis Function (RBF) as a classifier.

- Feature Extraction based on the importance of the word using Singular Value Decomposition (SVD)