CHAPTER 5
DISCUSSION, CONCLUSION AND FURTHER EXTENSION

5.1 Discussion

This thesis develops children gender classification system using speech. The task of gender identification from children's speech is difficult compared to adult gender identification. This system can classifies voice of children between 6 to 11 years of age with average accuracy of 83%. Based on the nonlinear nature of the data, classifiers efficient in discriminating non-linear data namely, RF, ANN, SVM, LR and GNB are considered. The RF outperforms the other classifiers with an average accuracy of 83% for gender classification. Classifiers can identify well for age group 9-11 years and most of miss classifications occur in age group 6-8 years.

5.2 Conclusion

This thesis presents implementation of children gender classification using speech. In the proposed system, voice feature extraction, machine learning (ML) and classification algorithms are combined. The system conducts a number of experiments with MFCC features dataset with the goal of developing a children's gender recognition system. The gender classification system is implemented by applying Python programming language and the experimental results have been analyzed. The best percent of divided the training data and testing data in this system is 70:30. RF classifier achieves the better classification results than other classifiers by using this test size. For train-test split testing, testing accuracy of RF model is 84%. RF classifier achieves 83 % for male prediction and 86 % for female prediction. For cross validation testing, the average accuracy of RF classifier is 83%. The analysis of the results shows that the performance of the proposed system is good, as the average accuracy of RF classifier is 83%.

5.3 Further Extension

For future work, the main goal is to increase the accuracy of gender recognition system. It can be extended to other researchers and can also be tested by using different features and other classification techniques. The performance of a few more different classifications would be compared to obtain a conclusion and the classification method with the best performance will be used in the speaker recognition system. Moreover, a perfect speaker recognition system would be developed with many functions. Then the research work will be focused on emotion recognition and region recognition. A voice recognition system can be designed which might be applied in Internet of things, such as user authentication by voice, where voice input will be used.

5.4 Limitation

The system is tested by using wave files with duration 1 second; most of wrong predictions is seen in predictions of the gender of children (6-8 years). Age group of 9 to 11 years can generally make correct classifications more than children of 6 to 8 years. The overall average accuracy is 83% with RF classifier.