CHAPTER 8

CONCLUSION AND FUTURE WORK

This chapter concludes the thesis by summarizing the contributions and findings of this research work and presenting the thoughts for future research in this area.

8.1 CONCLUSION

This research is concerned with the study and analysis of automatic cancer detection system to improve the efficiency and effectiveness of the solution for diagnosing cancer using data mining on bone marrow haematology data. The proposed system offers many advantages including better accuracy and greater automation. This system has been implemented using VB.NET and SQL in windows environment. The developed system is expected to provide valuable support to the physicians in decision making at the time of diagnosis.

In this research work three schemes have been proposed to improve the performance of CAD system to detect cancer. It includes a technique for association rule generation using various kinds of attribute selection, a technique for missing value imputation using associations between the attributes and a technique for data discretization. The improvements have been measured in terms of diagnostic accuracy in three techniques and the accuracy of filling missing values in the missing value imputation technique. The diseases considered in this work are blood cancer types Acute Myeloid Leukemia and Acute Lymphoblastic Leukemia. The data is collected from a pathology department of a reputed hospital and is used for performance evaluation.
8.2 THESIS CONTRIBUTIONS

The bone marrow reports are analysed and the relationships between the bone marrow attributes and the doctor’s final impression about the corresponding patient are identified. Various data discretization techniques are designed to discretize the continuous data of the bone marrow. The missing values are filled with various Missing Value Imputation techniques and are compared and the best result is used to fill the data and the filled data is used for better prediction of cancers. The major contributions to the thesis are summarized as follows:

- Identify efficient discretization techniques for preprocessing the haematology data
- Identify relationships among bone marrow attributes using attribute association rule generation
- Fill the missing values in the database using associations generated
- Better prediction of cancers using the discretization techniques, association rule mining and missing value imputation

8.3 FUTURE WORK

This research work can be further extended to diagnose any disease from the physicians report by extracting the diagnosis details and the associated factors for the diagnosing of diseases. The associations are generated using low, high and normal status of attributes of the data and they can be further divided into more than 3 ranges to find the associations that may improve the efficiency of the system.