CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENTS

This chapter concludes the thesis by summing up the outcomes of the research carried out, on preprocessed cluster oriented data reduction method based on considering any kind of data set in a secure preprocessed approach. The outcomes of the research are summarized, and possible future enhancements of this work are suggested.

6.1 CONCLUSION

The presented research work for identification of better data reduction in a large data set in data mining gives a new preprocessed cluster oriented data reduction method, providing prominence to the selection of better clusters or groups that best represents data reduction for the entire large data set. The method utilizes the preprocessing filters from both unsupervised and supervised and locates the best filter among them. Finally, the preprocessed data set are executed using various clustering and classification to find better groups or clusters to represent the data reduction. The proposed method and the results obtained by implementing the developed system the results have been obtained which have been further compared with the selection of attributes which also implements preprocessing filters. The existing clustering algorithms K-Means, EM and similarly, Meta and Tree classifiers performance are shown in terms of total time take to build the model in (secs) i.e. time complexities involved.
The results obtained here, have demonstrated that the proposed system can accomplish the goal of most efficient and optimum accuracy in data mining applications. The proposed method has exhibited the better data reduction in large data set in case of K-Means with Normalize filter in Instance level and when using the combination of selection of attributes with K-Means the Normalize filter again performs better than the Randomize filter which is a newly identified result respectively. On the other hand, in classification the Meta classifier using the random subspace performs better than the tree classifier methods. When comparing the clustering and classification results together the K-Means with Normalize filter with or without implementation of selection of attributes proves with better data reduction. The presented technique can play a vital role in reducing the time complexity in the mining operation. The results obtained have been illustrated and established as the best solution for reducing time complexity.

6.2 FUTURE ENHANCEMENTS

This research describes an efficient novel cluster forming in terms of data reduction data mining method and has proved the efficiency of the preprocessed oriented data reduction method. It is a fact, that the present research has demonstrated the best solution for achieving the goal of novel data reduction in the form cluster in data mining. The future of the work would concentrate on comparing the efficiency of the Preprocessed Oriented data reduction method with that of self organizing maps and in spatial time series dataset each referencing a location in common spatial framework the filter and refine algorithm can be proposed instead of correlation analysis whose cost involved is very high.

As the developed system architecture has accomplished all the desired and proposed goals for data reduction in large data sets in the form of
cluster formation, least time complexity and processing speed in response as the data set size is reduced. The comparative study of the existing preprocessing filters with K-Means and EM clustering algorithms has stated that the presented technique can be efficiently implemented and used for real time data mining applications.