## 7.4 Suggestions for further work

In this section we highlight some important aspects of the future work that can be done in this important area.

1. Current research has focused on some Machine Learning techniques to predict student potential and performance in technical programs. The use of Machine Learning techniques allows us to delete noisy data and find strong correlation patterns between selected attributes. Many other ML techniques and their hybrids can be tested on the datasets, so that more accurate results may be got.
2. The dataset used in this study was collected from different universities and institutions of Jaipur, Rajasthan only. This limitation can be removed and data collected from many other districts, not only in Rajasthan but in neighbouring states to get a general idea about students in the wider North-West region. Thus, many other educational Boards may be added in the dataset.
3. We have considered only those attributes of students which are normally used by educational Institutions, in the dataset. It is thought that many other attributes, like age of student, his/her mental level, in terms of IQ (Intelligent Quotient) or EQ (Emotional Quotient) may also be used to give a more accurate prediction of their potential.
4. Again, the size of the dataset is an important factor in such statistical analysis. Needless to say, more student data should be included in the analysis in order to improve the accuracy of the ML techniques used.
5. We have collected student data using the conventional means of physically moving to different Universities and educational Institutions and giving their students a Google Form to fill. This is a slow process. It is important to include new ways of collecting student information, taking advantage of social media and other on-line communication tools, like LMS, VLE and Wikis. These have low missing values and will thus speed up the data collection process.
6. Finally, this research has been restricted to using only Machine Learning techniques, like Decision Tree, KNN, Naïve Bayes, SVM, since the data sets used were limited. Large datasets, collected as given in the above suggestions, will allow many Deep Learning techniques to be used to predict student potential and performance at much higher accuracies. This will make it possible to build larger, knowledge-based Recommender Systems for use by higher educational Institutions.