Chapter 6

Conclusion

People are concerned about the education system and possible ways to improve education quality, especially in India. The main objective of any educational institution is to impart quality education and improve the performance of students’ for better managerial decisions. The way to reach the highest level of quality in an educational setting is the prediction of student’s success in a learning institution. Various prediction models are available to predict the performance of students’ based on different attributes. Although, various predictors are available for prediction there is no certainty observed in discovering whether a student will be an intellectual brilliance, a drop-out or a standard performer. Other predictors may be initially appealing, but some analysis shows that it is not the key method and other options are preferable.

Existing mechanisms have just not been able to respond to the emerging demands being made by professional education. This thesis intends to identify the most influential factors that affect the academic performance of a student and can help in predicting the at-risk students so that they can be corrected within time and hence improve their result. This study will benefit all the stakeholders in educational settings where early detection of at-risk students was not possible till date. This study will help in improving the overall results of the students by providing early detection of students’ at-risk so that they improve/correct the factors that are affecting the academic performance.

This study provides a prediction tool for the faculty members to identify the at-risk students and provide them with corrective and preventive actions that will result in improving the performance of the students which will ultimately help the faculty members to attain the best results for development in their professional career. The institutes will be benefited as more number of students will shift from high-risk category to low-risk category hence minimizing the number of failures, dropouts and increasing the retention, revenue, job opportunities and ranking of an institute.
An investigation has been performed to depict the importance of data mining techniques and classification to predict the academic performance of students in their early years of graduation. After investigating the Data Mining Techniques, an Ensemble model evolves as a better mechanism for efficient prediction. Factors that may affect student performance in professional courses have been considered using Classification Data Mining Techniques. This is the motivation for the proposed research work for predicting the performance of undergraduate students using data mining techniques.

To evaluate the progress of the students enrolled in technical courses, a study has been conducted to examine various factors that influence students and their performance. The educational institutions use automated computer programs/tools developed with different technologies to handle the growing data in the educational field. With growing technologies, predictive technologies have started growing tremendously to handle the huge databases in the field of data mining. The predictive technologies have collected a lot of literature through academic research in data mining. Data Mining is based on the theory that the historical data holds hidden and previously unknown information that can be used for the extraction of useful information. Similarly, the prediction of academic performance of the students can be done using data mining and assist in decision making. Moreover, it also explores how to control the factors that affect the performance of students.

At any given time, it is very difficult to foresee the exact academic performance in early years of graduation. So, the main objective of this study was to find a method that could efficiently predict the academic performance of students in colleges and universities. To accomplish this, some data mining techniques were selected to create an Ensemble for the above-said purpose and sought to predict the performance of students. The observations made in this research work.

i. The outcome of the study helps both the students and mentors to take better decisions. Based on the analysis, the students can improve on their grades taking care of a few parameters that affect their studies. This
analysis of variations can be used by mentors to make some remedial measures for the admitted students. It would likely to improve their academic performance, assisting them in studies, if required, and ultimately render improved results.

ii. The prediction of the academic performance of students is successfully achieved using the tool Rapid Miner studio 9.0.1. After the analysis, it is concluded that the student’s performance in their early years of graduation can be accurately predicted with the Ensemble Model. The student’s data is analyzed to find out the difficulties faced by students as well as mentors.

iii. Predicting students’ academic performance is a matter of concern in the early years of graduation using historical data of students. The dataset used in the experiment have the attributes which are different in nature: numerical, categorical and range based. They consist of marks of the previous semester, attendance of the present semester, study hours spent after college, assignment marks, and extra-curricular activities. Our Experimental results on the different set of values show that the classification algorithms perform more accurately. The proposed Ensemble is meant to find the factors that influence the academic performance of students.

iv. The success of an institute is mainly dependent on the student success rate. We have derived a prediction mechanism for the success of student by considering studying habits and grades. This method is proved to be effective due to the error percentage of classification calculated being less than 2%. By considering a variety of factors like demographics, academic, social and economic, additional processing can be performed. However, the method helps the mentors to improve their teaching methods according to students’ requirements.

v. Our data mining approach in predicting students’ academic results based on the influencing factors as evaluated in the previous chapter was to be tested for its efficiency. Validation is performed on the student’s data using corrective action and preventive action shown in the previous
chapter. Our Ensemble model provides better results as compared to the individual classification produce. However, the Ensemble model helps the students improve their performance and academic activities in order to improve their academic performance in the further year.

Considering the future of the country, it is imperative to groom the well-educated youth. Hence several training activities, and proposing an appropriate mechanism, plays a major role in taking our education system to the next levels. To come out with better results, educational data needs to be processed so that better predicting techniques can be found.

Turning to the conclusion, to improve student academic performance in addition to student retention, academic institutions and universities need to work on major influential factors. Attributes like family issues, previous academic data and socio-behavioural factors were taken into account as they play a prominent role in students’ scholastic performance. In this study, distinct attributes related to students’ academics, behaviour, demographics and social issues were considered. Student information is collected in their initial years and most influencing factors were included in the study. An Ensemble was formulated to investigate the academic performance of the students’ because the results produced with the help of Ensemble were highly accurate. Any kind of data related to education can be fed into the proposed Ensemble model and undoubtedly it will produce the same level of efficiency in results. The proposed Ensemble is a blend of Support Vector Machine, Naïve Bayes and Decision Tree as these three classification techniques prove to be more efficient as compared to other techniques.

Results obtained through Ensemble can be helpful to various stakeholders and can be considered for decision making. Moreover, these results will assist mentors to plan remedial strategies to help students’ in their initial years to score well in their exams. The Ensemble Model will assist the educational institutions and mentors to supervise the learning ability of the students and can keep track of their performance in an effective and efficient manner. Finally, this model can guide students’ to take preventive measures at an early stage so that their performance should not be affected further by factors considered in the study. This Ensemble
will come to the aid of stakeholders, mentors and most importantly students; it will also improve the learning processes to cut down dropout and failure rates.

As suggested in Chapter 5, implementing CAPA will be beneficial for the mentors as well as students to improve on weak areas of students. Also if there are some unavoidable circumstances faced by a student that leads to failure or lower performance grades then CAPA can help them to take timely precautions. Mentors can save their students’ being drop out after implementing CAPA and can provide them the best suggestive measures.

Students found at-risk after applying the data mining techniques were 173. Out of these 173 students 10 students were considered to take the validation test of the study i.e. CAPA. On the basis of the influencing factors, suggestions were made for the improvement of the performance of students in their academic career. Necessary preventive measures were suggested to students according to their requirement of assistance. Variations in their results were noticed after incorporating the suggestions provided in CAPA form.

To sum up, it appears that Ensemble model proves to be an efficient model for the prediction of academic performance of students at an early stage of their graduation (as early detection and prevention which can only be possible through the means of Ensemble). However, prediction through Ensemble model would appear to be a robust alternative.