Chapter 1

Introduction

1.1 Preface

The Education system in India has grown tremendously in the last few decades. Earlier it was confined to the four walls, with the students simply memorizing what was taught to them. The teacher-student relationship largely appeared more of a master-slave than of a guru who would understand students’ problems and deal with them at a more moralistic level. The students, too, on the other hand, felt obliged to the guru’s sayings. The students although, would be staying far from home but felt comfortable in being vocal about their problems to the mentor. With the passage of time and changes in the laws relating to the education ambience, the relationship developed to an ethical level; the teachers and the students equally felt a responsibility towards their jobs. Forlornly with the changes in the social setting and liberalism seeping in the society, the adolescence became more casual in their approach and acquired various bad habits that started to hamper their growth. This also led to a massive gap not only in the teacher-student relationship but in the parent-child as well.

With the growing need of technology, the setting of an Indian education system has evolved with new methods of education. It can be stated that it is the need of the hour to adopt advanced methods to improve the quality of education. Various schemes by the government such as accreditation of the schools, Pratham Scheme that focuses on the education of the underprivileged children, National Common Minimum Program (NCMP), The Central Advisory Board of Education (CABE-1935) to monitor educational programs, NCERT, Sarv Shiksha Abhiyan et al have helped reframe the educational setting of India.

Under such schemes, the government ensures that the teachers are trained to prioritize their focus on the students’, focusing not only on their moral or social growth but also the psychological. The schemes also target enhancing the level of education by inculcating technology. Such new technological apparatus has been
commendable, fetched and called for by the governments to provide a world-class education to the children. Additionally, the fair development in providing students healthy and hearty meals, smart classrooms, counselling sessions, one-on-one talks with the parents, has tremendously helped in providing a healthy atmosphere to a child to grow.

Despite the efforts made by the government, institutions, teachers and the parents, the Indian education system seem to encounter various challenges[109,110,111]. Researchers have further gone deeper to solve educational issues such as drop-out rate, failure and low rate of percentage of students. Figure 1.1 shows the challenges faced by the education system in India.

![Figure 1.1: Challenges in Indian Education System](image-url)
The globalized world has provided a mammoth dais to the education and learning to both the teachers and the students. Establishment of private institutions and increased competition amongst institutions forced education settings to modify their purpose and scope that will help them in the future as well. This has rendered a course of action that is commendable. The authorities have advertently framed a course of action for an arrangement of quality education and role of a teacher in the academic carrier of students. Since times immemorial, the role of teacher and the relationship between the teacher and student has been enormously received. Due to these guidelines, every university is eager to find out diligent pupils that will score good grades and add a extra feather in their caps and put institute in good positions. Such improvements in the education system will also create a level for education on which the future generations will grow. It will also render the education system on the tally at the international level.

Early detection and prevention of poor performing students can help at-risk students in the long run. Due to early detection, students and mentors both can work on the weak areas to improve the academic performance in the future. The mentors need to take necessary and important steps to achieve the success of students in their studies. Thus, the aim of mentors and students is to enhance the value of education in diverse educational settings. It is imperative to focus on the educational background so as to give a balanced platform to the students to grow. Henceforth, good results then will not only influence a students’ future but also help improve the status of the stakeholders.

However, there is the number of reasons that a students’ performance require more attention than only improving the teaching process. Early prediction of academic performance could be used as an indicator for their future results, such as students have to repeat the course if not scoring well, an assignment could be given to work on the weak areas. It is important for the teacher to plan necessary actions based on students’ results. The teacher could plan remedial actions, plan extra hours or assignment for the student to help overcome their weaknesses and offer concluding sessions. Besides this, the student that requires extra attention could be motivated to learn. The students who score better could prepare for competitive exams based on their better performance abilities.
Academic performance evaluation is done in the form of mid-term exams in the college and universities and special tests are taken to improve their results. Despite all these actions, special assignments and quiz and other learning games should be performed in the classes on a routine basis so that the required remedial actions can be performed on the students who need special attention and care to improve their academic performance.

Socrates too believed that the purpose of knowledge was to help people improve their lives. The mission of educational institutes is to create, collect and contribute to knowledge in a competitive environment. Generating knowledge inside the educational institutes is believed to be important to condense the boundaries of teaching learning process can be enhanced by applying the data mining techniques on the data set to get better results. Such techniques are applied on educational data to produce better results, for example, Prediction, Classification, Regression, Machine Learning, Neural Network, Association Rule Mining and Text Mining. Data mining techniques also render a new perspective to the students’ academic performance.

For the purpose of interrogating the teacher-student relationship and students performance, educational institutes are collecting a large amount of data referring to students. For an organization and management of the educational process the researcher read and re-read the data gathered in a calculative manner. However, available data is used to generate reports and easy to understandable queries that are rarely reaching needy people at the It has often been seen that much of the data remains unused due to the inability of educational administration to handle it because of large volumes and increased complexity. The learning algorithms can effectively transform available data into information and knowledge to support early detection and prevention. Early detection of the pre-requisites thus applied leads to a clear and informative data.

In the field of education and learning the students play an important role as they are the subjects who define a teacher’s performance as well as prepare a wholesome foundation for the future. Evaluating the performance of a student-at-risk and quantifying it to the next level to achieve the improved academic
performance of the student is of utmost importance. Performance measures are thus incorporated and furnish the researcher with information about students’ academic performance and also their ability to achieve predefined goals.

Factors that affect the performance of students can be categorized into four dimensions:

- **ACADEMIC**: Relating to performance and Assessment
- **SOCIAL**: Relating to society and moralistic growth.
- **DEMOGRAPHIC**: Relating to an area (Urban and Rural)
- **BEHAVIORAL**: Relating to the peer pressures, expectations of the parents and students’ psychology.

These attributes have often been widely used in various disciplines such as sports, psychology, sociology et al but will be used by the researcher to give a new dimension to the field of information technology.

The focus of the researcher was on the collection of the student data which comprised of various factors such as marks from the previous class, attendance, socio-economic factors, personal habits, the habit of studying and the time devoted to studies after college hours. These attributes illustrate information about students’ academic history and helped to develop a framework that will assist to improve the quality of educational settings in India.

**1.2 Problem Statement**

Student performance is one of the utmost importance when it comes to analyzing the problematic areas related to their growth. Their performance and analysis have attracted a lot of attention from the perspective of research. Increased rate of student failure and drop out in different educational streams has enforced significant pressure on the educational system to handle this issue carefully. Besides the government and the Ministry of Education taking the issue seriously, many researchers have re-considered the drop-out rate so as to provide alternatives to the raging problem.
The study has applied various performance measures on students’ data. They can be categorized into the following categories based on their academic performance:

- **High Risk**
- **Low Risk**
- **Medium Risk**

The students who score below 60% and have lower attendance falls into the category of High-risk and need special attention and assistance to improve their results. Medium-risk students are those who score between 60%-70% of marks and may or may not require special measures to improve their scores. Low-risk students are good performers with 80% and above also they do not need any assistance; they are already motivated to score well. The representation of categories of students is depicted in the form of a bar chart in Figure 1.2.

The focus of this study is early detection and prevention of high-risk students’ using data mining techniques to predict the academic performance of students in their early years of graduation and to study the influential attributes that affect the student performance in academics. The research used various learning algorithms which have not been explored to the extent necessary. Data mining is set on the theory that the historical data retains the hidden and unknown information that can be used to predict the future and can assist in important prevention measures.
It is, however, noticeable that with the advent of new technologies educational institutes are getting automated and data generation has increased tremendously. Predictive technologies have contributed a lot to educational research on the basis of this voluminous data, as it provides various ways to improve the educational settings. Preventive strategies are also taken into account in this study as they can help build a better future in education. The preventive measures can also render perspectives for further research to the high-risk students.

1.3 Motivation

Motivation is a strong factor when it comes to analyzing the issues relating to a topic. It facilitates the researcher to gather an in-depth analysis of the subject taken focusing especially on the variables that affect its existence. It further identifies the forces behind behavior—here; it is taken to be the factors that influence student performance. Motivation thus refers to a selection of qualitative research methods that probe the underlying behavior. It further helps raise questions that would answer the problematic areas of the subject undertaken.

The proposed work will focus on early detection and prevention of academic performance of the students in the Indian educational setting. For this, the researcher focused on large data sets of students which are classified according to their performance in the early years of graduation. As the academic record plays an important role in the prediction of students’ academic performance, it was taken as a prime variable in establishing the research project. Predicting the performance of a student in their first year of graduation can help them improve their results in the preceding semester of their professional course in higher education institutions. It is a challenging area for the researcher to understand different factors and behavior of students that affect their academic performance. This study will discuss the most relevant studies conducted in the field of education.

Many factors have led to the varying performance of the students. Following are the factors that influence the performance of students:

i. Type of classrooms i.e. traditional and online.
ii. Marks in previous semesters, class attendance, behavior towards assignment submission, socio-economic factors, educational background of the family, the income of the family, their awareness about opportunities provided by the institutes near them.

iii. Students’ academic, extra-curricular and assessment history in previous semester.

1.4 Research Questions

Based on the above-mentioned factors (that influence) the performance of students following questions will be answered:

RQ 1  What are the influencing factors that affect the academic performance of the student in graduation?

RQ 2  What will be the measures to detect and prevent the students’ at risk?

RQ 3  What will be the corrective and preventive measures that help students improve their results?

This study provides suitable solutions to the research questions mentioned above. The study elaborates the different ways of evaluating the performance of students in academics by taking care of the influential factors that put a negative or positive impact on their academic performance. After results are obtained from the study, corrective and preventive measures will be provided to the mentors to help students’ improve their academic career. The study will also influence further research that may be conducted in the field of academics.

1.5 Scope of the Study

The scope is a part of planning that involves documentation, goals, tasks, costs inculcated and deadlines. The researcher here planned to organize and interrogate a students’ academic performance – as a stakeholder as well and considered other stakeholders in the education field. The scope of the study was to answer the raging questions that a student faces, their poor performance and the remedies to deal with the crisis if one does occur. It additionally, focuses on the well-being of social order as students become the pillars of a nation.
Education is facing one of the major problems these days. Student’s attrition rate is considered as in the early years of their graduation they face difficulties in learning in a new environment. Post-school the students are fragile and under peer pressure that distracts them from their chosen path. Additionally, this vulnerable age and the raging hormones affect their studies; there is a need for adventure as there is number of distractions in their way, the number of factors affecting their academic performance. The factors include attendance, socio-economic factors, personal habit of studying, educational background of the family, the income of the family, and their awareness about opportunities provided by the institutes nearby. These factors help to take important decisions about student academic performance like the right time of intervention, the material that can help them to improve their academic performance.

To find out the best-performing students has been the focus and interest of many teachers and researchers due to a large percentage of failures in education. The emerging EDM has received considerable attention in dealing with uncertainties in today’s fast-changing educational environment, yet little empirical research exists to explain how the Indian education system can be improved. So, there is a need for a prediction model that can be applied to the Indian education system to achieve the information about superior performing students.

i. Analysis of the mechanism required for the improvement in the performance of the student at risk of failure.

ii. An analysis of the causes, consequences and existing problems faced by the student during the learning process.

iii. A framework is proposed to improve the learning process of a student who is at the risk of failure.

iv. Measures will be taken to cover the gap between the students who are at risk than who are performing well.

v. Mentors are requested to follow the teaching patterns diligently and cope up with the changing behaviour structure of the students.

vi. Institutes are asked to focus on providing better facilities to the students and follow the regulations and schemes as and when instructed by the government.
1.6 Thesis Organization

The organization of the thesis is as follows:

Chapter 2

The chapter elaborates exhaustive Literature Review and background of Data Mining Techniques, history of Data Mining Techniques in higher education and its applications in various other related fields. The literature review reflects the use of various data mining techniques in Higher Education Institutes for prediction, analysis, text mining, pattern matching etc. The study also gave an insight into the work predicted by scholars from various disciplines. The focus of this research will be mainly on predicting the performance of students’ in higher education based on the various identified factors from the literature review.

Chapter 3

The chapter illustrates the detailed Research Methodology that was adopted to achieve the objectives. The study can be broadly divided into two major portions i.e.

i) Data collection and pre-processing

ii) Development of prediction model

The study performed its predictions based on the collected data using data mining techniques. It will introduce the proposed ensemble model of data mining techniques for the prediction of academic performance of the students. The Chapter will also discuss in detail the Ensemble Model and its various variants.

Chapter 4

The chapter elaborates proposed work comprising of application of various Data Mining Techniques on the collected data and creation of Ensemble model based on best-performing techniques. This chapter concludes the important factors influencing the academic performance of the students. These factors were further used in preparing the Corrective
Action and Preventive (CAPA) as discussed in Chapter 5. CAPA has proven to be a reliable measure in consideration of analyses of the performance of the students.

Chapter 5

The chapter illustrates results obtained after applying the Data Mining Techniques on the data set. Influential factors extracted from high performing data mining techniques were used to develop a Corrective Action and Preventive (CAPA) Action form. CAPA will be implemented on the students that are found at-risk. CAPA is designed to suggest the actions based on the questions that are extracted through influential factors that affect the performance of the students’. The Chapter will also discuss the stakeholder’s and their importance in making a strong educational ambience.

Chapter 6

The chapter concludes the proposed work including the constraints and possible improvements as it relates to the performance of the students in future. With the results and analysis done, the researcher provides a clear picture of the pros and cons of the academic structure present in India. This chapter also elaborates the future scope of the study. It equally justifies the models and equations used by the researcher during the course of study.

The Reference section includes a bibliography that helped to conduct this study. The Appendices have the questionnaire and details of publications related to this research work.