

# Student Performance Prediction System

Sohan Kamble  
Student  
JSPM's Rajarshi Shahu College  
of engineering, Tathwade  
Computer Department

Prof.J.P.Kshirsagar  
Assistant Professor  
JSPM's Rajarshi Shahu College  
of engineering, Tathwade  
Computer Department

Kodam Saurabh  
Student  
JSPM's Rajarshi Shahu College  
of engineering, Tathwade  
Computer Department

Maske Adesh  
Student  
JSPM's Rajarshi Shahu College  
of engineering, Tathwade  
Computer Department

**Abstract** - The system aims at increasing the success graph of students using Naïve Bayesian and the system which maintains all student admission details, course details, subject details, student marks details, attendance details, etc. It takes student's academic history as input and gives students' upcoming performances on the basis of semester. The aim of the system to predict the student performance on the basis of different parameters related to previous year using data mining classification and clustering. An educational institution needs to have an approximate prior knowledge of enrolled students to predict their performance in future. This helps to identify promising students and also provides them an opportunity to pay attention to and improve those who would probably get lower grades. As a solution, we will develop a system too predict the performance of students from their previous performances using data mining classification and clustering. By applying the K-Means and its improved K-Means. The student performance is usually stored in student management system, in different formats such as files, document, records, images and other formats. These available students' data could be extracted to produce useful information. However, the increasing amount of students' data becomes hard to be analyzing by using traditional statistic techniques and database management tools. Thus, a tool is necessary for universities to extract the useful information. This useful information could be used to predict the student performance.

**Keywords**- Classification, Clustering, K-Means, improved K- means.

## 1. INTRODUCTION

The ability too monitor the progress of student's academic performance is a critical issue to the academic community of higher learning. Analyzing the past performance of admitted student's would provide a better perspective of the probable academic performance academic performance of student in the future. This can very well be achieved using the concept of data mining. We will propose system with better accuracy for predicting student performance using data mining classification algorithms. System will also help the weaker student to improve and bring out betterment in the result. Knowledge discovery in academic institution becomes more critical and crucial in terms of identifying the students' performance. In the extraction of actionable knowledge from a large database the data mining plays a vital role. The actionable knowledge extraction provides an interestingness and meaning to the mined data. This paper focuses on the prediction of the students' academic performance from the large student database. The mining algorithm like clustering and classification algorithm is revisited to predict the performance after initial mining of raw data. The main scope of this paper is to reveal the outcome of the performance analysis of a student .This work will help the university to reach betterment in providing the quality input to the student community and impart the knowledge effectively.

**SCOPE OF THE PROJECT**

Predicts the weak area of students and try to improve that.

The teachers can check the students overall performance through this system.

**2. PROPOSED SYSTEM**

There are few features from the existing systems that are employed during the design and implementation phase of the proposed system.

These features and functionalities include the user interface, students’ performance prediction, illustration displays and report generation. Good user interface provides a user-friendly interface as It is easy to be navigating and not complicated. Meanwhile, the students’ performance prediction is included into the proposed system to make sure the objectives are achieved. Furthermore, the generation of reports in Portable Document Format (PDF) and illustration display such as charts in PDF makes student performance analysis easier. From these features found in proposed system, all the user requirements would be fulfilled. Our Proposed system will provide following features:

- i. Able to help lecturers to automatically predict students’ performance in course “3rd year engineering”.
- ii. Able to keep track and retrieve students’ performance in a particular course and semester
- iii. Able to view the factors that affect the students’ Prediction result
- iv. Able to generate students’ reports.

The proposed system architecture is designed as shown in below:

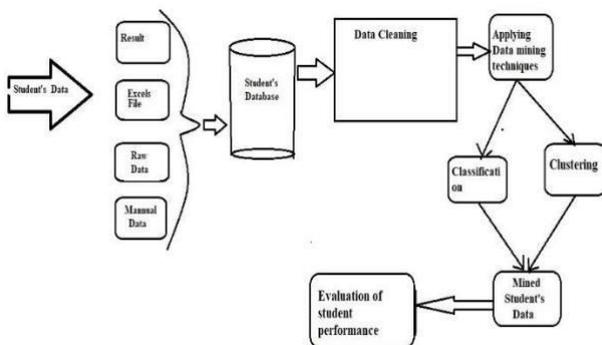


Fig. System Architecture

Given a student information file which is uploaded by teacher, our system will predict the performance of student mentioned in the file using data mining classification algorithms result will be displayed.

**2.1 PROPOSED SYSTEM FEATURES**

- 1. Data set is given to the system through the uploaded file.
- 2. System will apply classifications algorithm.
- 3. Student will be classified into pass or fail class.
- 4. Result will be displayed.

**2.2 IMPLEMENTATION OF PROPOSED SYSTEM**

We had divided the entire implementation into Four stages:-

- 1. In the first stage, information about students who have been admitted to the third year was collected. This included the details submitted to the college at the time of enrolment.
- 2. In the second stage, extra information was removed from the collected data and the relevant information was fed into a database.
- 3. In the next stage, the test data, i.e. information about students currently enrolled in the third year, was applied to the decision trees.
- 4. The final stage consisted of displaying of results.

**2.3 DATABASE**

We will take a training dataset consisting of information about students in third year Computer Department. This data was in the form of a spreadsheet and had details of each student such as full name, enrollment ID, gender, percentage of marks obtained in second year percentage of marks obtained in the entrance examination, admission type, etc. For ease of performing data mining operations, the data was filled into a database. Database also consists of list of the teachers associated with department.

**3. IMPLEMENTATION**

The implementation work is based on the collected data which possess various data mining aspects. The Student data is taken into account for the performance prediction.

The proposed research work is categorized into various modules. This research work is carried out with the inclusion of data mining technique and implementation software.

The proposed work speculated as the useful application where, student's performance can be viewed and placement criteria for various concerns are listed efficiently. The student performance is generated as a report graph which can be used as a survey to improve the student performance in the future. The student's and other user can login with their given credentials onto the desired application which modeled to predict the student performance and their placement criteria. The students, faculty and other user can view the past and recent performance of the respective courses to reexamine their current performance.

### 3.1. K-MEANS

**Step 1:** Accept the number of clusters to group data into and the dataset to cluster as input values

**Step 2:** Initialize the first K clusters

- Take first k instances or
- Take Random sampling of k elements

**Step 3:** Calculate the arithmetic means of each cluster formed in the dataset.

**Step 4:** K-means assigns each record in the dataset to only one of the initial clusters

- Each record is assigned to the nearest cluster using admeasure of distance (e.g. Euclidean distance).

**Step 5:** K-means re-assigns each record in the dataset to the most similar cluster and re-calculates the arithmetic Mean of the clusters in the dataset.

### 3.2. IMPROVED K-MEANS

The original k-means algorithm is modified to improve the accuracy and reduce execution time.

**Step 1:** Input: In this step take input from the user the dataset and pass it to the algorithm.

**Step 2:** Apply the normalization technique to the given dataset

**Step 3:** Apply the sorting technique to the given dataset

**Step 4:** Apply the algorithm to find initial centered from the dataset

**Step 5:** Assign data object to the centroids (repeat until convergence occur)

**Step 6:** Recalculate centroid

**Step 7:** Check for the convergence

In the modified algorithm first initial cluster size is calculated by using total attributes divide by number of

cluster. Then normalization technique is used to normalize the dataset to scale up the values is the range. In the

next step the sorting technique is used on the dataset because processing the sorted array is faster than unsorted array. Then calculate the initial centroid by mean of the cluster. In the next step assign the data object to the initial centroid by calculating Euclidean distance. Check for the convergence criteria. Repeat the steps until no more changes in the last centroids and updated centroid. Because of the initial centroid is generated by calculation the number of iterations is fixed, the initial centroids are determined systematically so as to produce clusters with better accuracy.

## CONCLUSION

In conclusion, the project 'Student Performance Prediction System' it describes the overall performance of the student ,in academic, sports , practical's, culture, social as well as extra curricular activities. his project help to select the field in future as well as the student weak in which area how to improve his knowledge in that area.

## REFERENCES

- [1] Chew Li Sa, Dayang Hanani bt.Dahlia Hos- sain, Mohammad bin Hossain, "Student Performance Analysis System (SPAS)".
- [2] Sagardeep Roy, Anchal Garg, "Analysing Performance of Students by Using Data Mining Techniques".
- [3] Suhas S Athani, Sharath A Kodali, Mayur N Banavasi, P. G. Sunitha Hire- math, "Student Academic Performance and Social Behavior Predictor using Data Mining Techniques"

[4] V. Shanmugarajeshwari,R. Lawrance , ”Analysis of Students’ Performance Evaluation using Classification Techniques”).

[5] Ms. Tismy Devasial, Ms.Vinushree T P2, Mr.Vinayak Hegde3, ”Prediction of Students Performance using Educational Data Mining”).

[6] Apurva Bandelwar1 Rutuja Athawale2 Rohini Dharmadhikari3 Prof. J. P. Kshirisagar4 “Student Portal with cloud Compiler”).