

# Student Performance Analysis System using Data Mining

Disha Kalambe

Department of Information Technology  
Atharva College of Engineering  
Mumbai, India

Surabhi Khedekar

Department of Information Technology  
Atharva College of Engineering  
Mumbai, India

Anita Labade

Department of Information Technology  
Atharva College of Engineering  
Mumbai, India

Komal Mahajan

Assistant Professor  
Department of Information Technology  
Atharva College of Engineering  
Mumbai, India

**Abstract**—In this age of computerization, education has also re-constructed itself and is not limited to old lecture method. Nowadays, lots of data is collected in educational databases, but it remains unutilized. In order to get required benefits from such a big data, powerful tools are required. Data mining is an emerging powerful tool for analysis. The previous system doesn't give the guidance to student based on the overall performance. The proposed system presents the analysis of student performance on the basis of academic performance, extra-curricular activities, strengths, weakness and hobbies. The proposed system uses classification algorithm and guides them by displaying the areas where they need to do improvement in order to contribute to a student's overall growth by generating a score card for the same. The proposed system will give all the required information of each and every student.

**Keywords**—Data Mining, Classification, Academic Performance, Score card

## I. INTRODUCTION

The academic performance of student is usually stored in various formats like files, documents, records etc. The available data would be analyzed to extract useful information. It becomes difficult to analyze student data by applying statistical techniques or other traditional database management tools. Hence there is a need to develop an automated tool for student performance analysis that would analyze student performance and will guide them by displaying the areas where they need improvement, in order to contribute to a student's overall growth by generating a score card for the same. The proposed system will display results of student performance on a single click action by the user, thus inducing automation and reducing efforts of staff in analyzing student performance manually. The proposed system finds out student trends on the basis of outcomes of students' academic performance, strengths, weakness, hobbies and extra-curricular activities. Academic data includes unit test, student's theory, practical's and term work marks. This data gathered will be processed by classification algorithm of data mining. A result from classification algorithm will be recognizing as Trend. This trend will help us to track where

the students excel and where not and what are their abilities which can be enhanced. The analysis will summarize the outcome and will classify students based on the results. This system will recognize interest of student in particular areas. The proposed system will also classify the students who are eligible for placements based on company's criteria.

## II. LITERATURE REVIEWED

Brijesh Kumar Baradwaj and Saurabh Pal have proposed a system that uses data mining methodologies to study students' performance in the courses. Data mining provides many tasks that could be used to study the student performance. In this research, the categorization task is used to evaluate students' performance and as there are many approaches that are used for data classification, the decision tree method is used here. Data like Attendance, Assignment marks Class test, and Seminar were gathered from the student's management system, to predict the performance at the end of the semester. This paper investigates the correctness of Decision tree techniques for predicting student performance.

Koma IJirage, Rucha Choudhari, Varsha Gupta, Anil Pathe and Disha Maind have proposed a system that is going to find out student trends on the basis of outcomes of tangible records as well as indirect surveys. This data collected will be processed by k-means technique of data mining. A result from k-means algorithm will be recognizing as Trend. This trend will help us to track where the students excel and where not and what are their abilities which can be enhanced. The analysis will summarize the result and will classify students based on the outcome. This system will recognize interest of student in particular areas. To empower our student for substantial contribution to technological, economical, social and entrepreneur progress of the society.

Bo Guo, Rui Zhang, GuangXu, ChuangmingShi and Li Yang have proposed a system that present a deep learning architecture for predicting students' performance, which takes advantages of unlabelled data by automatically learning multiple levels of representation. This system pre-train hidden layer of features layer wisely using sparse auto-encoder, and then use supervised training for fine-tuning the

parameters. This system train model on a relatively large real world students dataset, and the experimental results show the effectiveness of the propose method.

### III. PROPOSED SYSTEM

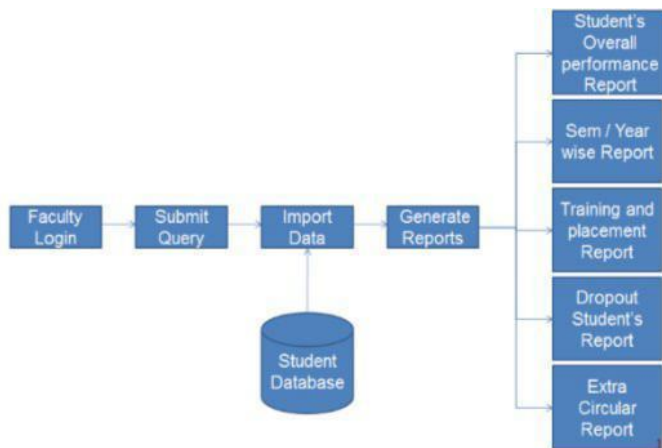


Fig. 1. System Block Diagram

#### A. Description of block Diagram:

The proposed system include an user profile creator to provide user interface, user login, student performance analyzer, Extra-curricular activities, Academic performance, score card generator, student performance credit card, student development card, achieved credit, passing criteria card and semester-wise student performance attribute card. Students will be grouped into categories like distinction, first class, higher second class, second class and pass class based on passing criteria.

- 1) *Admin Login:* This module is used for loginfor the admin user of the system, so that admin can manage staff, create and manipulate subjects and student, assign subject to each semester and staff to each subject.
- 2) *Faculty Login:* This module is used for loginto a faculty so that they can access the system and can add marks to each student for their respected subject. Also they can add strength and weakness of a student which can be both used to perform an analysis of a student.

*Student Login:* This module is used for loginthe student who can then view their marks for each subject entered by the staff.

#### B. Flow of the purposed system:

*Step1:* In the purposed system the user will first loginto the system to initiate the process.

*Step2:* Then the user will get a form which containsfields like Student Name, Student ID, Branch of the student, Academic Year, Semester ETC.

*Step3:* Then from each of the field the user will get adata of an student individual or Semester wise or Branch wise or Year wise.

*Step4:* Then it will also contain extra activitiesperformed by the students and from all that analysis the faculty will come

to know that the student is weak or strong and the student is eligible for the placements.

### IV. METHODOLOGY

#### A. WEKA:

It contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to these functions. The system is not so much a single program as a collection of inter-dependent programs bound together by a common user interface. Typically these modules fall into three categories: data set processing, machine learning schemes, and output processing.



Fig. 2. Steps of extracting knowledge

The proposed system extracts all the students' data from the database. Then selects the desired data, split the dataset into test and training sets and preprocess it as filter out features. Then the data is transformed into the data mining and the evaluation is done.

#### B. ID3 Algorithm:

ID3 is a simple decision tree learning algorithm.

The basic idea of ID3 algorithm is to build the decision tree by employing a top-down, greedy search through the given sets to test each attribute at every tree node. In order to pick the attribute that is most useful for categorizing a given sets, we introduce a metric - information gain. [3]

ID3 (Examples, Target\_Attribute, Attributes)

- 1) Create a root node for the tree
- 2) If all examples are positive, Return the single-node tree Root, with label = +.
- 3) If all examples are negative, Return the single-node tree Root, with label = -.
- 4) If number of predicting attributes is empty, then Return the single node tree Root with label = most common value of the target attribute in the examples.
- 5) Otherwise Begin
  - a) A = The Attribute that best classifies examples.
  - b) Decision Tree attribute for Root = A.
  - c) For each possible value,  $v_i$ , of A,
  - d) Add a new tree branch below Root, corresponding to the test  $A = v_i$ .
  - e) Let  $Examples(v_i)$  be the subset of examples that have the value  $v_i$  for A
  - f) If  $Examples(v_i)$  is empty
  - g) Then below this new branch add a leaf node with label = most common target values in the examples

- h) Else below this new branch add the sub tree  
ID3 (Examples( $v_i$ ), Target\_Attribute,  
Attributes – {A})
- 6) End
- 7) Return Root.

## V. CONCLUSION

The proposed system displays the results of student on a single click action by the user, thus inducing automation and reducing efforts of staff in analyzing student performance manually. The data collected is processed by classification algorithm of data mining. The proposed system helps to the students and the teachers to improve the division of the student. The proposed system also work to identify those students who require special attention to lessen fail ratio and taking proper action for the next semester examination. The proposed system also classifies the students who are eligible for placements based on company's criteria.

## REFERENCES

- [1] Brijesh Kumar Baradwaj and Saurabh Pal "Mining Educational Data to Analyze Students' Performance" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol.2, No.6, 2011.
- [2] Komal Jirage<sup>1</sup>, Rucha Choudhari, Varsha Gupta, Anil Pathe<sup>4</sup> and Disha Maind " Student Performance Analysis Based on Learning" International journal for research in emerging science and technology, volume-2, issue-3, march-2015. E-ISSN No: 2349-7610
- [3] Bo Guo, Rui Zhang, Guang Xu, Chuangming Shi and Li Yang "Predicting Students Performance in Educational Data Mining" International Symposium on Educational Technology, 2015
- [4] Snehal Kekane, Dipika Khairnar, Rohini Patil, Prof. S.R. Vispute and Prof. N. Gawande "International Journal of Innovative Research in Computer and Communication Engineering" Vol. 4, Issue 1, January 2016, ISSN(Online): 2320-9801, ISSN (Print): 2320-9798.
- [5] Pooja Thakar, Anil Mehta, Ph.D Manisha, Ph.D "Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue" International Journal of Computer Applications (0975 – 8887) Volume 110–No.15, January 2015.