Results and Placement Analysis and Prediction using Data Mining and Dashboard

Siddhi Parekh
B.E. Student,
Dept. of Information Technology
St. Francis Institute of Technology Mumbai, India

Ankit Parekh
B.E. Student,
Dept. of Information Technology
St. Francis Institute of Technology Mumbai, India

Ameya Nadkarni
B.E. Student,
Dept. of Information Technology
St. Francis Institute of Technology Mumbai, India

Riya Mehta
B.E. Student,
Dept. of Information Technology
St. Francis Institute of Technology Mumbai, India

ABSTRACT
In educational institutes a huge amount of data is being generated. The produced data does not provide enough information which obscures important details of data that may help in better understanding of available data and its utility. If the data is analyzed efficiently, it can provide many insights, specific information regarding various facets of data which can be useful in a multiple ways. Analysis of data plays a very important role in understanding of information from a given set of data. Analysis of data can be performed using various data mining algorithms which help them to take decisions or arrive at a conclusion with the help of available data. This paper limits implementation of an efficient decision making system which will enable the college to analyze the results and placement of its institute. The main objective of this system is to generate query specific reports of the academic performance of a group of students or a student in particular which helps in evaluating student's potential strengths and weakness with respect requirements of various companies for placement, which assists in understanding of placement trends. Dashboard representation provides a platform to prospect the overall performance of the system.

General Terms
Data Mining, Dashboard, Result and Placement Analysis

Keywords
Classification, Data Mining, Result and Placement Analysis, ID3, Prediction.

1. INTRODUCTION
Existing technology existing in many Indian universities does not use integrated systems, creating inconsistency in data. Hence the available information is neither reliable nor consistent and it is difficult for the Educational institute to interpret their progress. It is inconvenient to track the trends or patterns of the results and the placements, making it more difficult to change the policies or any other change required for improvement. This happens due to varying data formats or faulty data entry or improper data updating. Therefore propose of web based application is to enable the storage and visual interpretation of data on the same platform, using integrated system and allowing a proper and broader analysis providing valuable information that is dependable.

Analysis and prediction often becomes an arduous task in absence of the right mining techniques and technology. Prediction of students' performance is very important in academic environments. This project implements such diverse mining functionalities and various set of algorithms to provide an evaluative interpretation of the students' data and provide clear view of the academic performance.

Hence Data Mining serves as a technology which enables various such organizations and institutions in determining the progress and generates a report for the same. It is a solution to various tedious processes involved in determining and calculating the statistical calculations and helps save time.

2. EXISTING SYSTEM
In present day scenario where education has been privatized and where there is ceaseless competition amongst students and institutes, there is a need to be more organized and have the ability to make sound decisions and make constructive changes. Due to these challenges, the management is meeting the diverse needs of utilizing a decision support system for facing increased complexity in academic processes and college policies to keep their institution at a respectable position in the contest. This system helps in continuous improvements in operational strategies based on accurate, timely and consistent information.

- The data is maintained in different departments in distributed locations.
- Management finds it difficult to locate the reports needed by them as the data is in segregated form.
- The user interface for the current system is not satisfactory and in no way helps the decision makers.
- When the consolidated report from two or more different subject area is required, it is almost impossible.
- Even after collection of data from different departments, the data format may be non-uniform and will require structuring.

3. PROPOSED SOLUTION
As mentioned above the problems with non-consistent data formats and individual reporting, they aim to build a system that is aimed at developing a web based application for the analysis of the students' academic performance and placement results of the institute. It is an online application that can be accessed throughout the organization and prospective students or staff and others. It helps them to manage student information like marks obtained by students in various semesters, subjects, attendance and admission information greatly simplify and speed up the result preparation and management process. They can take micro-level decisions in

a timely manner without the need to depend on their IT staff. They can perform extensive analysis of stored data to provide answers to the exhaustive queries to the administration team.

4. SYSTEM DESCRIPTION

The proposed system is user friendly web based application developed for educational institutes that uses a single centralized integrated database for storing all information regarding students and other college activities like training and placements. All the information available in the storage is not made accessible to every user. Each user will be provided with a user name and password for authentication and restricted access and rights to information to protect the integrity and availability of information. It creates query specific reports according to user requests which can be observed in a visually enhanced manner using bar graphs or pie charts etc. on the dashboard for better understanding of the information displayed. The systems prediction module actually evaluates the student's ability and potential using and relating various parameters (such as academic performance, internship and other student information.) and provides with suggestions on improving the quality of students helping the institute to take decision on replacing or modifying their coaching , placement training or other management and admission policies. The system composed of the following modules.

5. SYSTEM MODULES

5.1 Centralized Storage and Data Mining

The system is consists a single integrated storage which is used to store myriad of data from all the different departments like admission, results and placements in a single common data format, avoiding inconsistencies and reducing paper work. Science of Data Mining helps them in extracting hidden knowledge from the database and they are using the same for our system. Implementation data mining functionalities they can identify patterns, clusters and data sets of a particular group and generate reports. Steps involved before starting the mining process are as follows.

- Picking the data points that need to be analyzed
- Data Collection in structured format.
- Data Cleaning
- Extracting the relevant information from the data
- Identifying the key values from the extracted data set
- Interpreting and reporting the results

5.2 Dashboard

A dashboard is a graphical representation of data which gives the whole overview of the current situation of an organization. The dashboard is easy to read or understand as it gives all the required data in a graphical format using various types of charts and graphs. Dashboard acts as a BI tool which enables easy overview of the database which provides valuable reports as per user specified constraints.

In the proposed project, the software will generate summary report regarding student information, semester wise marks list and performance report, which can be visually enhanced for better understanding, helping in prediction of future placement and academic results for decision making based on past records and correlations between students results and other important factors to improve students' academic performance and placement. All of the above project features will be displayed using dashboard.

5.3 Analysis

Analysis of data helps in extracting hidden knowledge from various forms of raw or inconsistent data which otherwise is considered to be futile or irrelevant.

Classification is a data mining function which gives sorts data into a class. The goal of classification is to predict what the target class of a particular case in the data is based on what it has learned from previous knowledge.

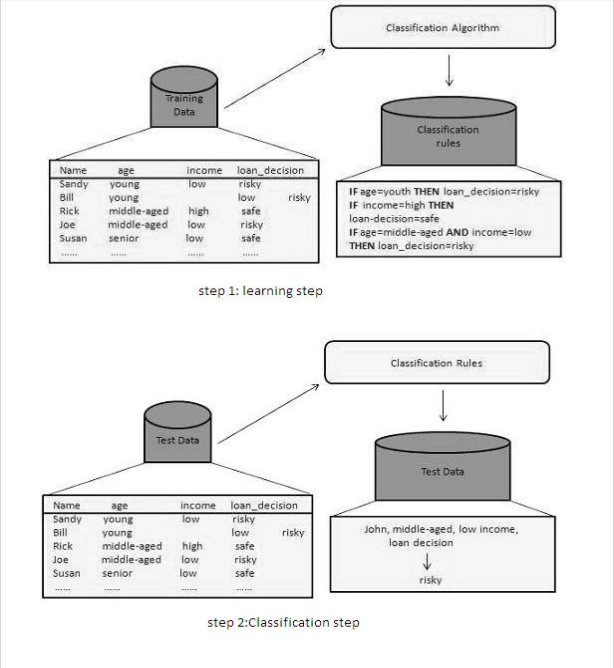

Fig 1: Classification Algorithm Steps.

As shown in fig. 1 the classification algorithm has 2 steps.

In the first step based on the learning data i.e. the data which is available it builds the decision tree or the classification model. In the second step the input data is classified according to the classification rules generated.

The prediction module of the system uses classification technique to predict the future trends of results and placements of students based on various parameters such as their previous academic performance. The classification model will be run on a selected set of data such that the required result will be used as a class label. The prediction module actually evaluates the he students' academic performance, also considers the various co-curricular activities of students in attempt to understand the various potential and lacking abilities of students required to be placed. It provides suggestion that helps the college teaching and placement staff in changing their teaching and training policies or adding various new additional programs which would help the student to develop the required skills.

5.4 ID3 (Iterative Dichotomiser)

Decision tree algorithm such as ID3, C4.5 can be implemented on structured database to generate reports. ID3 is an algorithm invented by Ross Quinlan used to generate a decision tree from a dataset. ID3 is the precursor to the C4.5 algorithm, and is typically used in the machine learning and natural language processing domains.

The best algorithm based on the collected placement data is ID3 with an accuracy of 95.33%. It works on the principle of the Occam's razor and used to create the smallest possible decision tree. It takes all the attributes which are unused and promotes the calculation of entropies which are used to measure the informative of node. It also scans and chooses the attribute which has the entropy is less or when information gain is large.

5.5 Implementation and API

The system will be implemented in the form of website. The website will be developed using PHP as the front end and MySQL as the backend database. To make the charts and graphs i.e. the dashboard, they are using Google Charts API. Google charts API is a free and customizable API provided by Google which has a rich gallery of charts and gives the ability to connect dynamic data with interactive controls and dashboards. On launching the web application, the user can request a query to the server for undertaking desired analysis. The server will send the requested query to the database to obtain results. On implementation of the algorithm for the query sent by the user, the results are sent back to the server. Similarly, various users can request the server via PHP forms to receive a response.

5.6 Academic Analysis

The given diagram snapshot shows the year wise analysis of overall pass percentage and performance of student. For example, in the academic year 2010-2011 the pass percentage of semester 6 students was greater than the pass percentage of semester 4 whereas in the academic year 2011-2012, the pass percent of semester 4 students was greater than the pass percent of semester 6 students. This will help them to create reports of students their performance whether it is increasing or degrading and necessary measures can be taken.

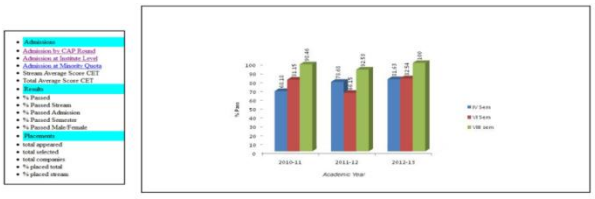

Fig 2: Pass percentage of students for 4, 6 and 8 semester.

Here, in the snapshot, they cater to a particular department and trace the performance over the future semesters. The pattern shows them that there has been an increase in the overall academic results which indicates the administration cadre to recognize the growth overtime.

5.7 Placement Analysis

Fig. 2 shows the detailed placement analysis which shows the number of students who have been placed over years. The given snapshot describes the total number of students placed by which companies they have been selected. Also, it shows stream wise distribution of placed students. For example, in all the academic years, computer science students have always excelled in placements followed by IT and lastly EXTC. The snapshot shows that there has been an inconsistent growth in the placement results where computers have always had the highest number of students placed and the lowest being EXTC. They cater to develop a system where 100% students are placed**.**

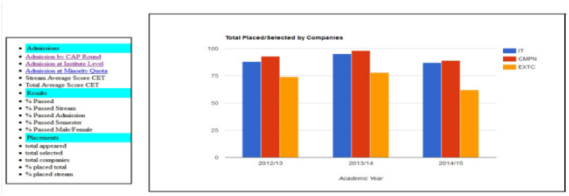

Fig 3: Total students placed of different departments in the years 2012/13, 2013/14 and 2014/15.

6. CONCLUSION

The system to be developed aims to provide an efficient single point management system which will give all the data of the students of the college at the same place. The system will provide a dashboard which will show the statistics in the form charts and graphs which will give easy interpretation of the academic state of students at any point of time. Generally the office staff members or the faculty of various colleges have to face issues when it comes to providing the upper management with reports of the state of academic and other activities related to the students. This is due the fact that all this information has to be managed department-wise and not centrally. Hence, the need to develop a system that can solve the above mentioned problem arises. This software comes with just that solution i.e. put the available voluminous data into a form of charts and graphs which is simple to understand. Thus it will give an immediate analysis of data and will help in the comparative study of data. The system as whole in not only beneficial to college but also to the students as it helps the college in understanding the potential of their students and the training that the college is required to provide their students in order to make the students technically and mentally ready for industries. On other hand it is very helpful to the students, as it aid them in realizing what potential opportunities await them if they select to take admission into a particular institute.

7. FUTURE SCOPE

Employing such a system is a proactive way to use data to manage, operate, and evaluate educational institute in a better way. Depending on the quality and implementation of the underlying data, such a system could address a wide range of problems by distilling data from any combination of education records maintenance system. It helps an educational institute improving the quality and placements of students being graduated from their institute. This system can also be implemented in different non educational institutes like business corporates, sports academies and manufacturing companies where the challenge would be taking into consideration the current market scenario as one of the most

important factors affecting quality of their products and employee.

8. REFERENCES

[1] S.SATHISH1, C.R.RENE ROBIN2 "An Optimized Result Analysis System for Institutions in India having Credit System" Vol 03, Issue 01; January-April 2012 ICICES-2012-SAEC

[2] Ajay Kumar Pal 2013 Classification Model of Prediction for Placement of Students, I.J.Modern Education and Computer Science

[3] Alex Berson, Data Warehousing Data Mining & OLAP, Computing Mcgraw-Hill, November 5, 1997.

[4] Arshad Khan, SAP and BW Data Warehousing, Khan Consulting and Publishing, LLC (January 1, 2005) Classification Model of Prediction for Placement of Student.

[5] Abel U. Osagie, Abu Mallam, 2014. " Students Record Analysis and Examination Result Computation Algorithm", International Journal of Technology Enhancements and Emerging Engineer Research, VOL 2, ISSUE 8 49 ISSN 2347-4289.

[6] T. Jeevalatha ,N. Ananthi, D. Saravana Kumar"Performance Analysis of Undergraduate Students Placement Selection using Decision Tree Algorithms" International Journal of Computer Applications (0975 – 8887) Volume 108 – No 15. December 2014.

[7] Jermy Smith, Robin Naylor," Determinants of degree performance in UK universities: a statistical analysis of the 1993 student cohort", Oxford Bulletin of Economics and Statistics, 63, 1 (2001) 0305-9049.