Prediction of Students Recruitment Process Using Data Mining Techniques with Classification Rules

T. Malathi[1], S. Srinivasan[2], K.R. DilliRani[3]

[1]M.Phil, Research Scholar, [2,3] Assistant Professor
[1,3] Department of Computer Science,   Dr. MGR. Chockalingam Arts College. Arni, Tamil Nadu, India
[2]Department of Computer Science and Engineering, Annai Mira College of Engineering and Technology, Vellore, Tamil Nadu, India

Abstract: **Data Mining is the process of analyzing data from different dimensions and summarizing it into useful information. Classification is a data mining techniques used to predict group memberships for data instances. In this paper, Classification is applied to analyze the Student's Recruitment process like Selected, Waiting and Not Selected. This paper will help the Lecturers to shortlist the students and arrange them for the recruitment process. Existing system of student recruitment process is to identifies the weak students before the final exam through internal assessment like, quiz, lab-performance, attendance and previous exam grade to improve the performance of the student in the final exam. Proposed system of student recruitment process is to apply the recruitment process based on the student's performance by taking appropriate steps at right time.**

Keywords: **Classification, predict, data mining, instances, membership, dimensions, shortlist.**

1.   INTRODUCTION

Data mining has a vital role in extracting hidden information from the given raw data, which is highly useful for researchers who make use of large data. The process of extracting such an effective, understandable data from the raw database is known as Data Mining process. This paper is mostly useful for the interviewers at the selection process, were the students data had been categorized by their grade systems to makes the filtering easily. This paper focuses by exploring the personal and professional data of all the students. These data will be highly helpful in short listing the students based on their rank, which is detected by applying rules such as clustering, association and classification rules of data mining.

Hidden pattern can be extracted using data mining from large database. Medical, marketing and real estate's are some applications of data mining. Data of educational field can be managed by a new data mining technique named educational data mining. Classification, association and clustering rule can be applied in decision tree, neural network and many important techniques of data mining to gain surplus knowledge. Student performance analysis can be done at the end of the semester exams by using the educational data mining techniques. Cluster analysis will be used in data mining to extract a reasonable data from the given sorted raw data. Clustering is a process of grouping highly similar data objects to make completely non similar groups in a given data. Data clustering can be used to extract positional useful data from large data-set. The usage of educational database has been increased rapidly. Clustering creates a homogeneous groups of data based on students ability, which enhances the quality of education. It also used to discover the relationship between the data-sets to help us make the decision efficiently. Clustering help us to gain the knowledge about the higher educational system. In this paper, clustering technique may be separated by two ways namely hierarchical and partitioning cluster respectively.

2.   PROPOSED SYSTEM

Academic performance calculated through internal and external assessment. Here, assignment, quiz, class test marks, lab performance are called as internal assessment. Previous exam grade and final exam grade called external assessment. To classify the previous exam grade by taking the internal assessment using data clustering techniques. Using classification technique to specified into topper, average and below average.

If prev_exam_grade=Topper,

Quiz=good work, assignment=completed, lab performance=perfect, Class-test=first mark, attendance=100% and then

final_exam_grade=topper.

Example:

If (mark<90) then Topper //eligible for all companies

If prev_exam_grade=average,

Quiz=good work, assignment=not completed, lab performance=perfect,Class-test=average mark, attendance=100% and then final_exam_grade=average.

Example:

If (mark>60) then Average//eligible for minority companies

If prev_exam_grade= below average,

Quiz=average work, assignment=not completed, lab performance=bad, Mid-term=low, attendance=irregular and then final_exam_grade=below average.

Example:

If (mark<60) then below average//ineligible

These are all the conditions help to classify the student performance and calculated by using data clustering techniques. This is the best way to improving the student performance for the purpose of student recruitment process.

In this paper, the proposed model forms the serious harm. That is, it is used to classify and identifies the weak students before the final year examination by appropriate steps. These steps are mostly used by teachers at right time to develop the student performance.

3.   DECISIONTREE ALGORITHM

A decision tree is a tree-like structure to represent the internal node to tests the attribute, each branch represents outcome of test and each leaf node represents class label. it has represented by classification rules.

Step 1:

Accept the number of nodes and the tree to group the tree into and the dataset to node as input values

Step 2:

Initialize the first node of the tree as root node.

- take first node instances or

-take random sampling of node elements

Step 3:

Calculate the arithmetic mean of each node formed in the tree

Step 4:

Decision tree assigns each record in the dataset to only one of the initial node.

-Each record should assigned as follows:

If the mark<90 then the result should be Topper its eligible for all companies,

 If the mark>60 then Average its eligible for minority companies and

 If the mark<60 then below average its called as ineligible

Step 5:

Decision tree reassigns each record in the tree to the most similar node and re-calculates the arithmetic mean of all the nodes in the tree.

4.   RESULTS AND DISCUSSIONS

Here, we groped the student in to three ways:

The possible grade assigns same as the possible labels.

The students can be grouped into three categories into topper, average and below average

class	CGPA	selected student	percentage of class
1	2.20-2.20	5	8.33
2	2.20-3.00	10	16.67
3	3.00-3.32	17	28.33
4	3.32-3.56	15	25
5	3.56-4.00	13	21.67

The categorized students should follows the two class labels "passed" for 2.20 grade and "failed" for 2.20 grade should be less than or equal to.

To classify the students among their CGPA that means from CGPA 2.20 we have 8.33% of students. From 2.20-2.30 student percentage is 16.67% from 3.00-3.32 we have 28.33%, 3.32-3.56 of student percentage is 25%. the student percentage of 21.67% between 3.56-4.00 CGPA. The graphical representation of CGPA and students percentage given below
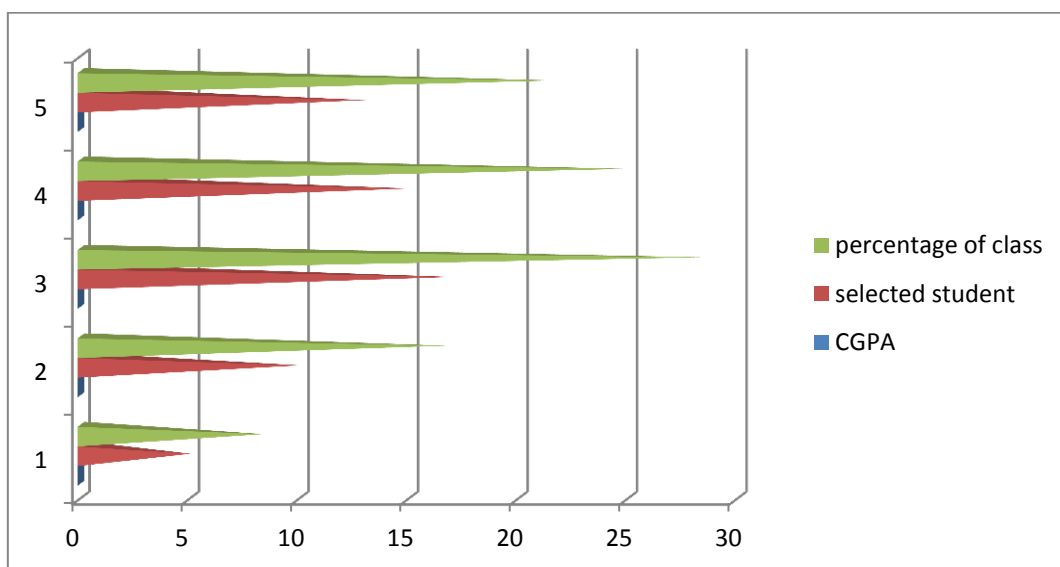
After classifying the student, we grouping the students into three categories, one is topper second is average and the last one is below average.

CLASS	CGPA	NO. OF STUDENT	PERCENTAGE
TOPPER	>=3.50	28	46.67
AVERAGE	2.20<=cgpa<3.5	27	45
BELOW AVERAGE	<=2.20	5	8.33

Graphical representation of student range is given below:
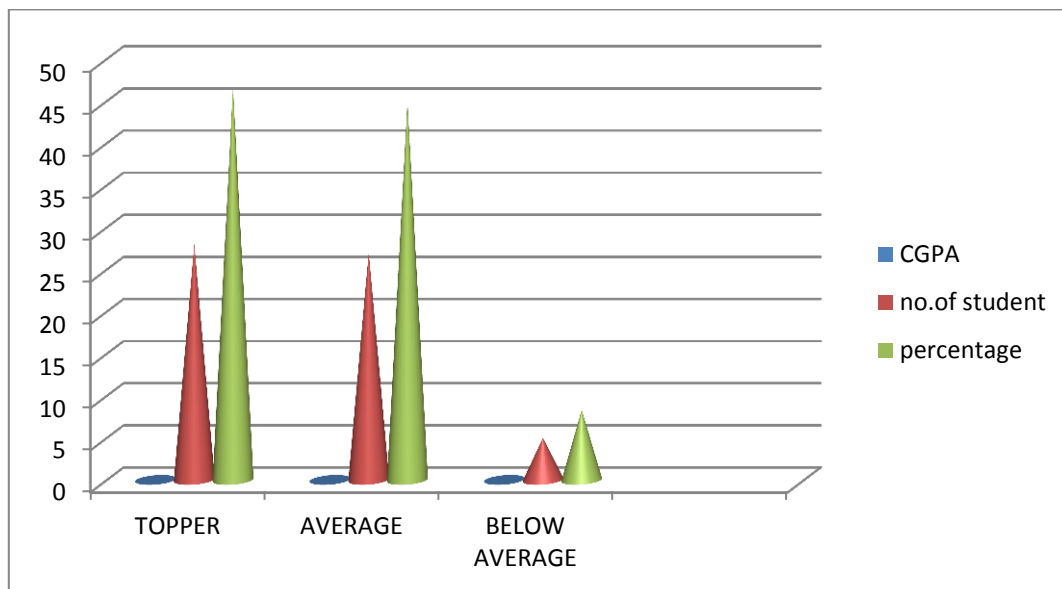
5.   CONCLUSION AND FUTURE ENHANCEMENT

In this paper level of students has been categorized efficiently by use of data mining process. Instructors can make use of the information generated after the implementation of data mining technique. One important function of data mining classification analyse to discover data sources distribution of information. Classification algorithm is used to group the students based on their proficiency. As a result it improves student's performance and reduces failing ration by taking appropriate actions to improve the training quality.

As a future enhancement, we use to fine-tune our technology to get more advance and accurate outcome. Finally will give some useful information to the instructors to improve the students learning ability.

REFERENCES

[1]   Md. Hedayetul Islam Shovon* and Mahfuza Haque(2012)Prediction of Student Academic Performance by an Application of K-Means Clustering Algorithm.

[2]   Umamaheswari K * and S. Niraimathi(2013)A Study on Student Data Analysis Using Data Mining Techniques