Classifying Students Performance by analyzing Internal Assessments of Student Data

Mr. M Satish Kumar[a], G Sasi Kumar [B] and H Nawaz [C]
[A] Assoc. Professor Dept Of MCA, SVCET Chittoor
[B] PG Scholar, SVCET Chittoor,
[C] PG Scholar, SVCET Chittoor

Abstract-Data Mining is used to extract meaningful information to develop relationships among variables. In this paper the student's performance are analyzed in order to learn how internal assessments will affect the final result. Here classifier algorithm is used to analyze the result attribute. The student internal assessments factors like mid terminal marks, assignments and attendance are studied. This analysis will help the faculty members to motivate students by predicting the final result.

Key-Words- Data Mining, classifier algorithm, Assessments, Predicting, Analyzed, Student Performance

I. INTRODUCTION

Data Mining is data analysis methodology used to identify hidden patterns in a large data set. It has been successfully used in different areas including the education environment. Data mining in educational field is a interesting research area which extracts useful, previously unknown patterns from educational database for better understanding improved educational performance and assessment of the students learning process. It is concerned with developing methods for exploring the unique types of data that confirm educationalenvironment which include student results repository.

As a students' repository is a collection of large amount of students' data, Data Mining and Classification technique can be applied to find interesting relationships between attributes of students

As a Fig.1 shows the concept of Data mining, this involves three steps:
Capturing and storing the data
Converting the raw data into information
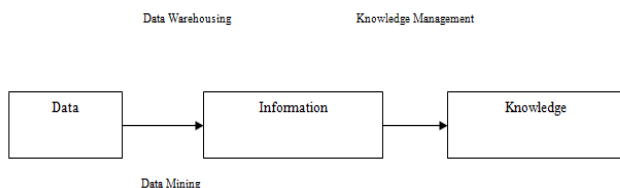Converting the information into knowledge

Fig. 1 Concept of Data Mining

Data in this context comprises all the raw material an institution collects via normal operation, capturing and storing the data is the first phase that is the process of

applying mathematical and statistical formulae to "mine" the data.

The collections of randomly selected student data are examined and assessed by the small group of teaching faculty members. An education reform for the 21st century has generated various models of learning of the performance of students. The assessment of student learning is an essential component in evaluating the overall institutional mid terminal marks, submission of assignments and attending to the class works regularly using J48 classifier algorithm.

Decision Tree analysis is a popular data mining technique that can be used to explain the interdependencies among different variables. In this paper decision Tree is used to extract the classification rules for the student data.

In this paper we analyze which internal assessment will affect more to get better results in the final university exams. It will help the academic planners in enhancing their decision making process in improving students performance.

The remaining sections of the paper are organized as follows: Section 2 describes the importance of Classification in analyzing. Section 3 explains about Data preparation for analyzing. Section 4 reports students' performance by analyzing internal assessments applied on data set. Finally we conclude this paper.

II. IMPORTANCE OF CLASSIFICATION

Databases are rich with hidden information that can be used for intelligent decision making. Classification is a kind of data analysis that can be used to extract models describing important data classes or to predict future data trends. This analysis can help us with better understanding of the data. Classification predicts categorical label.

A.Data Cleaning:-
This alludes to the preprocessing of information in place evacuate or diminish the clamor (by applying smoothing systems, for instance) and the treatment of missing qualities (e.g. by supplanting a missing worth with the most regularly happening quality for that characteristic, or with the most plausible quality in light of measurements). Albeit most arrangement calculations have some mechanisams for taking care of boisterous or missing information, this step

can help decrease disarray amid learning.

B.Relevance Analysis:-

A number of the traits in the information may be excess. Relationship examination can be utilized to distinguish whether any two given properties are measurably related. Case in point, a solid connection between A1 and A2 would recommend the one of the two could be expelled from further investigation. A database might likewise contain superfluous traits. Trait subset determination can be utilized as a part of these cases to discover a diminished arrangement of properties such that the subsequent likelihood circulation of the information classes is as close as would be prudent to the first appropriation acquired utilizing all characteristics. Subsequently, significance examination, as relationship investigation and trait subset choice, can be utilized to identify qualities that don't add to the arrangement or expectation undertaking. Counting such characteristics might somehow lull, and conceivably delude, the learning step.

Information change and lessening: The information may be changed by standardization, especially when neural systems or strategies including separation estimation are utilized as a part of the learning step. Standardization includes scaling all qualities for a given quality so they fall inside a little determined extent, for example, -1.0 to 1.0 or 0.0 to 1.0. In routines that utilization separation estimations, for instance, this would avoid characteristics with at first expansive reaches (like, say, wage) from out weighting qualities with at first littler extents, (for example, paired traits).

Information cleaning, pertinence examination (as connection investigation and trait subset determination), and information change are depicted in more prominent subtle element.

C.Comparing Classification and Prediction Methods:

Grouping and expectation techniques can be contrasted and assessed concurring with the accompanying criteria.

Exactness: The precision of a classifier alludes to the capacity of an offered classifier to effectively foresee the class mark of new or beforehand inconspicuous information (i.e., tuples without class name data). Additionally, the precision of indicator alludes to how well a given indicator can figure the estimation of the anticipated property for new or already inconspicuous information. Precision can be assessed utilizing one or more test sets that are free of the preparation set. Estimation systems, for example, cross-approval and bootstrapping, are depicted. Systems for enhancing the precision of a model are give on the grounds that the exactness registered is just an assessment of how well the classifier or indicator will do on new information tuples, certainty points of confinement can be processed to help gage this appraisal.

Speed: This alludes to the computational expenses included in producing and utilizing the given classifier or indicator.
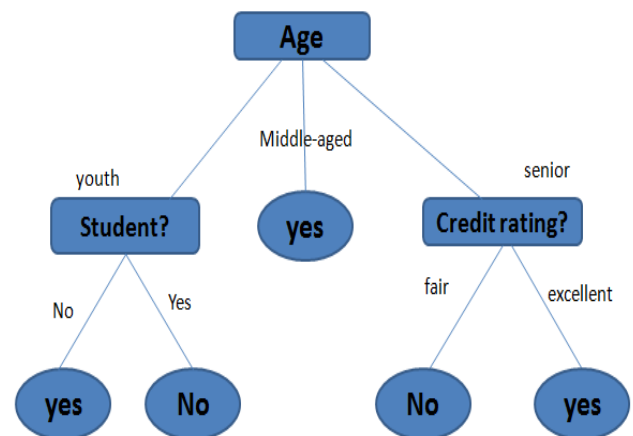
Vigor: This is the capacity of the classifier or indicator to make right expectations given boisterous information or information with missing qualities.

Versatility: This alludes to the capacity to build the classifier or indicator productively given a lot of information.

Interpretability: This alludes to the level of comprehension and understanding that is given by the classifier or indicator. Interpretability is subjective and in this manner more hard to evaluate. We talk about some work around there, for example, the extraction of grouping guidelines from a "discovery" neural system classifier got back to proliferation.

III.CLASSIFICATION BY DECISION TREE INDUCTION

Choice tree incitement is the taking in choice trees from class-marked preparing tuples. A choice tree is a flowchart-like tree structure, where every inward hub (non leaf hub) means a test on a quality, every branch speaks to a result of the test, and every leaf hub (or terminal hub) holds a class mark. The highest hub in a tree is the root hub.

"Why are choice tree classifiers so prevalent?" The development of choice tree classifiers does not require any space learning revelation. Choice trees can deal with high dimensional information. Their representation of procured learning in tree structure is instinctive and for the most part simple to acclimatize by people. The learning and grouping ventures of choice tree affectation are straightforward and quick. As a rule, choice tree classifiers have great exactness. In any case, effective utilization may rely on upon the current information. Choice tree incitement calculations have been utilized for order as a part of numerous application ranges, for example, drug, assembling and creation, budgetary investigation, cosmology, and atomic science. Choice trees are the premise of a few business standard instigation frameworks

Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCACI-2015 Conference Proceedings

A.   Categorization procedure

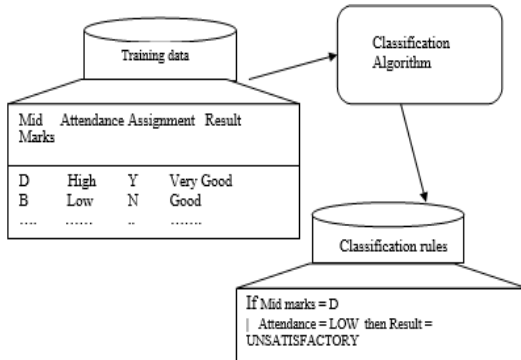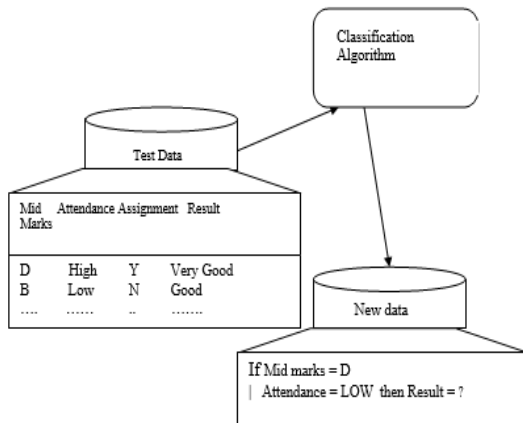Data classification is a two-step process, as shown in Fig. 2

(a)

(b)

Fig. 2 Classification Process

2.1.1 Sub-subsection

When including a sub-subsection you must use, for its heading, small letters, 11pt, left justified, bold, Times New Roman as here.

IV.TROUBLE CLARIFICATION

Figures and Tables should be numbered as follows: Fig.1, Fig.2, … etc Table 1, Table 2, ….etc.

If your paper deviates significantly from these specifications, our Publishing House may not be able to include your paper in the Proceedings. When citing references in the text of the abstract, type the corresponding number in square brackets as shown at the end of this sentence [1].

V. FINAL (CONCLUSION)

Please, follow our instructions faithfully, otherwise you have to resubmit your full paper. This will enable us to maintain uniformity in the conference proceedings as well as in the post-conference luxurious books by WSES Press. The better you look, the better we all look. Thank you for your cooperation and contribution. We are looking forward to seeing you at the Conference.

REFERENCES:

1.   X1. Author, Title of the Paper, International Journal of Science and Technology, Vol.X, No.X, 19XX, pp. XX-XX.
2.   X2. Author, Title of the Book, Publishing House, 19XX