Evaluating Students Performance in Placements Activity

S. Indhu Priya, PG Scholar, Department of Computer Science & Engineering, Coimbatore, India.

Dr. P. Devaki, Professor & Head, Department of Computer Science & Engineering, Coimbatore, India.

ABSTRACT: In an education system, predicting student’s performance in placement has become more challenging due to the large volume of data and imprecise data with fuzziness in educational databases. Large volume of data is processed using big data analytics methods. Processing of data with different factors and with different parameters is difficult in traditional system, where big data analytics can help organizations to better understand the information contained within the data. It also helps them to identify the data that is most important for the prediction and future decision making. The aim of evaluating student’s performance is to help them to develop individual student’s professionalism, to encourage self-improvement, to maintain achievements and also to give them prior idea about their level of skills in placements. It also plays a vital role in increasing placements. In this paper some of the existing methodologies and their drawback for the student analysis have been discussed.

Index Terms – Data mining, Big Data Analytics, Decision tree.

I. INTRODUCTION
Student’s performance is an essential part in higher education institutions. This is because the excellence of a high quality university is based on its academic achievements. It is achieved by showing better placement record. Students play a vital role in this. To achieve this student need to know their level of skill set and the area where they are strong and where they are weak, where they need to improve to place their level for the top companies. The current scenario of students getting placed is less when compared to the existing scenario, it’s because nowadays market is down due to the automation, and also service based companies are mainly based on US projects, but currently projects from US also being reduced this became the main reason for service based companies to dropout the employees. As from the NASSCOM report the intake for the service based companies will be reduced in the future years.

Product based companies are in prediction to hire candidates with high expectation but nowadays students in engineering are not fully capable to fulfil their expectation. Product based companies are expecting students with good percentage, and excellence in their programming languages with high ranking certification in top reputed organization. Were all these companies will not provide training for the students; they need to learn themselves and update all the current technologies.

II. PERFORMANCE PREDICTION METHODS:
DATA MINING:

In educational database the data are stored in a large number with different types of input, which are related to the performance of staff and students. Using data mining technique their level of performance is evaluated. Wanli, X., et al [1] synthesize learning analytics approaches, educational data mining (EDM) and HCI theory to explore the development of more usable prediction models. This method used to predict the students final performance in academics, they use the data from collaborative geometry problem solving environment. The approach described diminishes data dimensionality and systematically contextualizes data in a semantic background.

Yusof, N. A., & Fauzi, S. N. F. M., [2] analysed the students’ performance in practical training (PT) it provided a mining method that predict the student who is good at his academics will be able to perform good in his practical training also, with the same grade. The criteria used to assess are: students’ commitment, work quality, effort, communication
skills, comprehension of work, self-confidence, and technical skills. To assess this, 3 performance scales have been used. The data were analysed using descriptive analysis in SPSS.

ISSUES IN DATA MINING TECHNIQUE:

Learning model developed with the less performance and the data processed is not accurate. In data mining all the input are pre-processed which leads to the additional storage consumption and it all leads to two-step process. Learning model for academic is identified but with less performance.

FUZZY MODELING APPROACH:

The fuzzy method is used to handle the large number of vague and imprecise data that are of different format. Fuzzy models operate on information granules that are fuzzy sets and fuzzy relations. Information granules are abstract realizations of concepts used in modelling. As modelling is realized at higher, more abstract level; fuzzy models gives rise to a general architecture in which we highlight three main functional modules, that is input interface, processing module, output interface.

Neogi, A., et al[3] presents a mathematical model to evaluate faculty’s teaching performance using fuzzy logic. This method evaluates the degree of satisfaction from the data obtained from the expert group of people. The result is calculated based on all the topics in the subject. The obtained results from the proposed approach are compared with the conventional non-fuzzy approach and the comparative results are presented. As a result fuzzy approach is accurate than the non-fuzzy method. Arbaiy, N., & Suradi, Z.,[4] is also evaluated the staff appraisal using fuzzy method. The staff performance appraisal may involve judgments which is based on imprecise data especially when human (the superior) tries to interpret another human (his/her subordinate) performance. The performance appraisal system can be examined using Fuzzy Logic Approach and this was carried out in the study. The study utilized hierarchical fuzzy inference approach. The fuzzy inference has the human reasoning which uses approximate information and uncertainty to generate decisions. It consists of rules, facts and conclusions. The fuzzy inference can be implemented using the if-then statements or Fuzzy Associative Memory (FAM). The if-then implementation is the same as that executed in expert system except that it involves the linguistic variables [4].

The single inference structure is shown as Fig 2-a. Therefore, as more than one attribute of inputs are involved in this system, then a hierarchical inference structure is needed.
Such fuzzy rules are represented as given below.

IF work_achievement is 'poor' AND skill is 'poor' THEN appraisal_rate is 'need improvement'.

IF work_achievement is 'poor' AND skill is 'satisfactory' THEN appraisal_rate is 'need improvement'.

IF work_achievement is 'poor' AND skill is 'proficient' THEN appraisal_rate is 'meet expectation'.

The rules are followed and with the certain method an end result of appraisal are obtained for particular staffs that satisfy all the inference stated in the beginning of the fuzzy.

DECISION TREE:

Educational institutions are using educational data mining methods which is used to gain deep and thorough knowledge to enhance its assessment, evaluation, planning, and decision-making in its educational programs. EDM will help academic programs identify and discovered hidden patterns in the data. These extracted patterns can be used to predict student performance and behaviours easily. There are many techniques in data mining that can be applied to educational data, such as classification, clustering, and association rules to name a few. These techniques will help extract hidden knowledge and useful information [11].

The decision tree classification involves the steps before the implementation of algorithm were the data are collected in the proper order as they are programmed by the particular institution. Al-Barrak, M. A., & Al-Razgan, M.,[5]. Proposed decision trees for predicting students final GPA. In this paper, they used educational data mining to predict student’s final GPA based on their grades in previous courses. In their case study, they collected student’s transcript data that included their final GPA and their grades in all courses. After pre-processing the data, they applied the J48 decision tree algorithm to discover classification rules. The method extracted the most useful information that is used to predict the final GPA of a student using decision tree and also based on the most important course and their study plan for their grades to get higher. Guarin, C. E. L., et al[10] proposed the prediction model for low academic performance prediction using data mining techniques. The models aim to predict attrition in the student’s first four enrolments. First, considering any of these periods, and then, at a specific enrolment. Historical academic records and data from the admission process were used to train the models which were evaluated using cross-validation and previously unseen records from a full academic period. Romero, C., et al[7] proposed student’s final performance predicting from participation in online discussion forums. The method proposes the use of different data mining approaches for improving prediction of student’s final performance starting from participation indicators in quantitative, qualitative and social network forums. The forum plays an important role in collaborative learning in which two main actions are carried out by students: writing and reading. The students can adopt different attitudes in any of the actions, thus defining their behaviour profile such as active leaner, lurkers, etc.,

Fig3: Data mining approach

Fig3 provides the model for data mining were the data are gathered, pre-processed before predicting the rules and model, the prediction involves differ classification and clustering algorithm discussed above and an inference is obtained from the all the processed data.

Huang, S., & Fang, N.,[8] developed the predictive numerical models for predicting the student academic performance. This study develops and compares four types of mathematical models to predict student academic performance in engineering dynamics – a high enrolment, high impact, and core course that many engineering undergraduates are required to take. The four types of mathematical models include the multiple linear regression models, the multilayer perception network model, the radial basis function network model, and the support vector machine model.
ISSUES IN DECISION TREE:

The decision tree induction is able to handle the data set with entries in thousands only when large volume of data are need to be processed the data mining and decision tree induction method results in less performance and handling of those data become difficult.

EDUCATIONAL DATA MINING:

Educational Data Mining (EDM) refers the research field which is concerned with the applications that are developed by means of using data mining, machine learning and statistics. This EDM interests on the development of advanced tools and algorithms to realize the data patterns. EDM investigates the learning theories and updates the educational practice. Main objective of EDM is to predict the students’ future learning behaviour with the use of student modelling; this goal can be achieved by creating student models that incorporate the learner’s characteristics, including detailed information such as their knowledge, behaviours and motivation to learn. The user experience of the learner and their overall satisfaction with learning are also measured. Discovering or improving domain models through the various methods and applications of EDM, discovery of new and improvements to existing models are possible. Examples include illustrating the educational content to engage learners and determining optimal instructional sequences to support the student’s learning style.

PROCESS INVOLVED IN EDM:

The first phase of the EDM process is discovering relationships in data. This involves searching through a repository of data from an educational environment with the goal of finding consistent relationships between variables. Several algorithms for identifying such relationships have been utilized, including classification, regression, clustering, factor analysis, social network analysis, association rule mining, and sequential pattern mining. Discovered relationships must then be validated in order to avoid overfitting. Validated relationships are applied to make predictions about future events in the learning environment. Predictions are used to support decision-making processes and policy decisions.

TYPES OF DATA USED IN EDM:

The data gathered from online web-based educational learning systems and tools comes in different formats that are specific to that particular tool. It is certainly possible that the data generated from an educational system is also stored into a database. But these types of data are stored in the database; modern databases are capable of storing multiple formats including .dat, .mdl, .text, or log file itself. For example, an educational system may generate data in one of the following forms: a text format, web server log file, or LMS log files.

Handling of these different kinds of data using data mining technique is not as efficient. To handle these data in an efficient and accurately an approach of Big Data Analytics is incorporated. The Data analytics are capable of handling large volume of data and the data that are stored in different formats.

BIG DATA ANALYTICS OR LEARNING ANALYTICS

Big data analytics is the process of examining large data sets to uncover hidden patterns, unknown correlations, and other useful business information. Big data is a term for data sets that are so large or complex that traditional data processing applications are inadequate to deal with them. The term big data often refers, simply to the use of predictive analytics, user behaviour analytics, or certain other advanced data analytics methods that extract value from data, and seldom to a particular size of data set.

DIFFERENCE BETWEEN BIG DATA AND DATA MINING

Big data is a term of a large data set. Big data are those that outgrow the simple kind of database and data handling architectures that were used in earlier times, when big data was more expensive and less feasible. For example, sets of data that are too large to be easily handled in a Microsoft Excel spread sheet could be referred to as big data sets.

Data mining refers to the activity of going through big data sets to look for relevant or pertinent information. Data mining can involve the use of different kinds of software packages such as analytics tools. It can be automated, or it can be largely labour-intensive, where individual workers
send specific queries for information to an archive or database.

**BIG DATA ANALYTICS IN PERFORMANCE PERDICTION:**
Learning analytics is referred as the use of data, which may include big data, to provide actionable intelligence for learner and teachers. The current interest in learning analytics reflects wider interests in Big Data and Educational Data Mining (EDM). Big data has been described as an all-encompassing term for any collection of data sets so large and complex, that it becomes difficult to process them using traditional data processing applications whereas Educational Data Mining (EDM) describes a research field concerned with the application of data mining, machine learning and statistics to information generated from educational system.

**CONCLUSION**
The placement activity of institution is playing a vital role in placing their standards and names higher for the future college admissions. In this paper, a survey is made for college activities were students are needed to be focused to improve their placements and different ways to find the solution. Where many authors have proposed a method to improve the performance of staff but there is no related work for student’s placement activity. Future work will involve the implementation of evaluating students’ performance in placement activity and ways to improve their skills using fuzzy method.

**REFERENCES**