

Placement Analysis using Data Mining

M.Nesamanikandan, Dr.M.Kannan, R.Pushpanathan

Department of Computer Science, SCSVMV University Enathur, TamilNadu, India

Abstract— Education data mining is one of the growing fields of the present time as it grows many issues to improve system comes in the notice one of them is improvement of the placement. Placement is a very important issue for any educational organization. Every organization wants to improve its placement. Success of any educational institute is measured by the placed student of the organization. This paper actually deals with the application of neural network to the educational data to improve placement. In today's world all organization faces one of the big problems is recruit right candidate for the suitable position, Organization ready to invest a huge amount to recruitment process but till now they failed to recruit. In this paper, we apply the data mining techniques for placement prediction. To predict the performance of a student's is the great concern for the organization, as they seek knowledgeable, talented and qualified professionals to need to fill up their positions. According to the survey, the corporate companies spend a sum of 1800crores for choosing candidates to fill up their vacancies. The Majority of the companies recruiting the candidates via on-campus recruitment and to fill up them positions. Our method is very useful for corporate companies, consultancies. This method is the best way to get the right candidate at the right time in the corporate world. The industry gets the best talent candidates from different institutes/universities, and the students also get a chance to kick start their career with some of the best organization. But the students facing some difficulties in getting placements. To overcome the problem we apply the Improved Decision Tree classification algorithms on these data, we have predicted which students placed in Recruitment Drives. Corporate companies need only knowledgeable and skilled persons for the vacant position. To find that particularly skilled person there's a question that how can the companies identify them. In order to overcome this problem in this paper, we provide a complete solution to the recruitment process. Actual challenges appear when they will develop real-world software. Training develops confidence in whatever School, Colleges, Universities will train. But all corporate expect skilled, confidence with active persons. We are giving to find the right candidate for the right job in this development. After recruiting employees corporate feel much confident about their development. So as much

as possible, clear them confusions and get new ideas about their project training and become confident about their work.

Keywords— Educational, Data mining, predication, random tree.

I. INTRODUCTION

We The extensive scope of growth in data creates and initializes the various new techniques, algorithms, procedures and methods to access them reliably, to identify the hidden patterns and to extract the actionable knowledge. Since the size of the database is too large the data mining process is carried out with utmost concentration to reach a feasible solution.

Nowadays the data stored in the educational database increases rapidly. These databases contain hidden information for improvement of students' performance. This academic performance is influenced by many factors therefore, it is essential to develop predictive data mining model for students' performance so as to identify the difference between good learners and bad learners. In my point of view, a research methodology was native to generate a database. The raw data was processed in terms of filling up missing elements, transforming values in one form into another and relevant attribute/ variable selection. Many universities are trying to fetch the information about the student. Predicting the performance of the students is a challenging task. The main goal is to improve quality and standard of education in the university. This work also helps in growth of institution as well as the student. It helps to view both academic as well as personal details of the student which would be also helpful in evaluating student's overall performance. This work illustrates graphical view of the student performance, overall mark and it helps to check whether they are eligible for placement or not. Student evaluation has become a part of most institutions. This helps to provide a feedback for instructors to improve their teaching practices which would result in improved student performance. It also helps the administrator of the institutions to keep track of their institutions performance. As the resource pool grows, the information from students who have more talented with these resources also increases. The knowledge extraction techniques used as all data of student personal details to academic records in either in flat file or in excel form. The campus placement

of the students plays an important role. The companies identify the talented and qualified professionals before they completed their studies. So the main success of institution is giving the placement chance to the students. The main goal of this paper is to classify the on tree algorithms. Majority of students in higher education joining a campus oriented colleges. Therefore taking a wise career decision regarding the placement after completing a particular degree is good in a student's future. An educational institution contains a huge number of students records. Therefore searching a patterns and characteristics in this large amount of data is not a tough job. Higher Education is categorized into professional and non-professional education. Professional education provides more knowledge to professional students so that they can make their stand in private sector. Professional education may be technology oriented or it may be totally improving management skills of candidate. Masters in Computer Applications (MCA) course provides professional computer technological education to students. This course provides state of the art theoretical as well as practical knowledge related to IT and make students eligible to stand in running information industry. The prediction of MCA students where they can be placed after the completion of MCA degree will help you to improve knowledge of students future life. It will help to build reputation of institute in existing similar category institutes in the field of IT education. The present study concentrates on the placements prediction of MCA students. I applied data mining techniques using Decision tree and Naïve Bayes classifier to interpret potential and useful knowledge.

II. OBJECTIVE

The major objective of campus placement is to identify the talented and qualified professionals before they complete their education. This process reduces the time for an industry to pick the candidates according to their need. It is a cumbersome activity and hence majority of the companies find it difficult to trace the right talent. Many students do not understand the importance of placement training that is being imparted, whether it is aptitude training or soft skills. They show the least interest in this due to various factors viz., projects, assignments or more of activities loaded by the colleges as part of their curriculum thinking that it is not useful. It is the responsibility of the companies training on placement to make the students equipped on all aspects of career development along with creating a very good impact on them which makes them feel every minute they spend in the placement training session is worth being there and will help them in getting placed in their dream companies.

III. IMPLEMENTATION

3.1 Problem Statement

Choose a suitable student is a common need for any organization. So a data mining model has to be built which can predict the most suited student who supplies his information. Their performances are to be compared based on statistical measures. Also a generalized design framework for a typical placement prediction problem has to be formulated.

To predict the student's placement status, finding is not so easy task, because the placement depends on various factors. Sometimes a student who have good academic record and still not placed, on the other hand a student who has not good in his academic life he got placed. Sometimes it depends on luck and time and the frequency of companies which comes in college campus. It is hard to analyses using above factors to predict the student's placement. Initially, the paper has dataset of students' academic record; to analysis the pattern from given dataset, which affect the placement status. Secondly, to collect the large dataset which is also a difficult task? If the paper have large real dataset, then analysis can be more preciously resulted, which would be more accurate. Third thing is to collect a lot of information about student, which is called attributes, as the paper is not sure which attributes effect the output ,so if the paper have a lot of attributes, then the paper can find precious result ,which will be more accurate our model for prediction. Similarity measure is used to find the pattern in the given object. There are various mathematical methods, which is used to find the pattern from the given data set. Sum of difference is one of them.

3.2 Scope of the Study

The main aim of this research work is to identify relevant attributes based on academics, skills and curricular of final year student and design a model which can predict placement of the student using a classification technique based on decision tree. This model can be useful for faculties, university and students to more emphasize on those which are not eligible for placement according to this model.

Recruitment is one of the most important functions for any organization as they seek talented and qualified professionals to fill up their positions. Majority of the companies have been focusing on campus recruitment to fill up their positions. This method is the best way to get the right resources at the right time with minimal cost and within minimum time frame. While the industry get the best talent from different institutes/universities, students too get chance to start their career with some of the best companies in the corporate world at the beginning of their career. The focus of this paper is to identify those set of

students that are likely to face difficulty in getting the placements. The result of analysis will assist the academic planners to design the strategy to improve the performance of students that will help them in getting placed at the earliest.

The main objective of this paper is to use 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 classification task is used to evaluate student's performance and as there are many approaches that are used for data classification, the decision tree. Information's like graduation result, mca, result, hsc result, communication skills, programming skills, Details were collected from the existing database.

3.3 Proposed Systems

Prediction of student's learning attitude by using student models that integrate all correct and clear information of students it may be student's knowledge, behavior; academic performance, Class test, spoken English etc. The present day situation is that corporate companies are spending a huge investment for recruitment. If one can make filtering process of resumes easier and simpler then companies may not spend the huge amount depends upon the recruitment. Proposed model for the above problem is to classify the resume of the applicant based on some decision parameters. Decision Tree algorithm in the domain of Data Mining provides a significant road map for the classify the students. This paper is to identify those set of students that are likely to face difficult problem in getting the good placements. This process reduces the time for an industry to pick the candidate and train them according to their need while the students get exposed to the private sector environment at the right time and learn how to prepare themselves for the competition.

3.4 Advantages

To identify those set of students that are likely to face difficulty in getting the placements.

- Its reduces the investment cost
- Reduces time to recruit new employees

Easy to predict which students will place in Recruitment Drives.

IV. SUMMARY AND DISCUSSION

In this paper, based on MCA placement analysis and decision making using decision tree algorithm is being used to details for MCA department students of reputed engineering college. This methodology will assist the corporate companies and Placement In charge with Head of Departments in identifying set of skilled students that are likely to face problem during final placements. This

identification will help most of the corporate with students and college placements In charge to design the strategies to improve the academic result, improve the student's programming skills and communication skills. This model will play important role in improving the overall placements of the institute. This will result in significant improvement in quality of admissions and academic results in subsequent years. Like this institutes will be able to sustain their existence competing with the good Indian and Foreign institutes.

4.1 Future Enhancements

The present research will have been planned to extend by using neural network algorithms, genetic algorithms and support vector machines and to add aptitude skill and technical skill as attributes in the dataset. And also we will collect dataset of final year students till four semesters and applying various classification algorithms to predict the placement of students before final semester. In future we focused on online verification, job profile and conduct online test if both are matching , that students is well knowledgeable candidate for suitable positions, in our existing work student details given by college staff there is a possibility to prediction errors

4.1.1 Limitations

- There is chance of prediction errors
- Student maybe misbehave and given wrong information because college staff only giving student details.

V. SCREENSHOTS



Fig.5: Admin Login



Fig.5.1: Admin Login Page



Fig.5.2: Provider Reg.



Fig.5.3: View Details

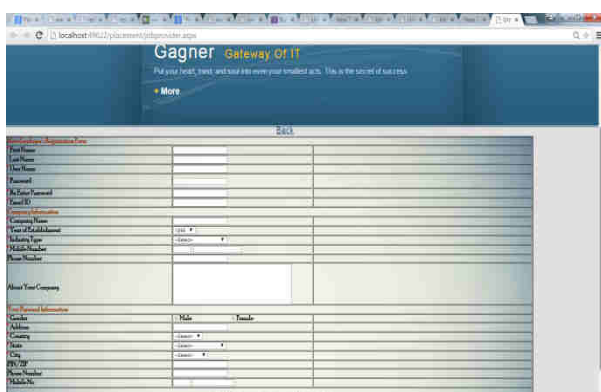


Fig.5.4: Admin Login Page

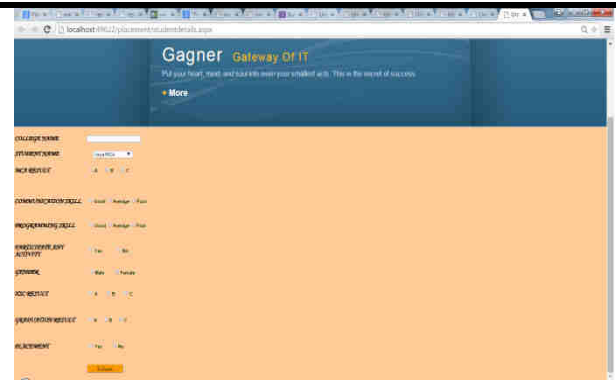


Fig 5.5: Student Register

VI. ACKNOWLEDGEMENTS

Firstly, we would like to express my sincere gratitude to my advisor Prof.Dr.M.Kannan for the continuous support of my M.Phil study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my M.Phil study.

Last but not the least, I would like to thank my family: my parents and to my brothers and sister for supporting me spiritually throughout writing this thesis and my my life in general.

REFERENCES

- [1] Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems.
- [2] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic (1996). "From Data Mining to Knowledge Discovery in Databases" (PDF). Retrieved 17 December 2008.
- [3] to a b c d "Data Mining Curriculum". ACM SIGKDD. 2006-04-30. Retrieved 2014-01-27.
- [4] Clifton, Christopher (2010). "Encyclopædia Britannica Definition of Data Mining". Retrieved 2010-12-09.
- [5] Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome (2009). "The Elements of Statistical Learning Data Mining, Inference, and Prediction". Retrieved 2012-08-07.
- [6] Han, Jiawei; Kamber, Micheline (2001). Data mining concepts and techniques. Morgan Kaufmann. p. 5. ISBN 9781558604896. Thus, data mining should have been more appropriately named "knowledge mining from data," which is unfortunately somewhat long
- [7] See e.g. OKAIRP 2005 Fall Conference, Arizona State University About.com Datamining
- [8] Witten, Ian H.; Frank, Eibe; Hall, Mark A. (30 January 2011). Data Mining Practical Machine

- Learning Tools and Techniques (3 ed.). Elsevier. ISBN 978-0-12-374856-0.
- [9] Bouckaert, Remco R.; Frank, Eibe; Hall, Mark A.; Holmes, Geoffrey; Pfahringer, Bernhard; Reutemann, Peter; Witten, Ian H. (2010). "WEKA Experiences with a Java open-source project". *Journal of Machine Learning Research* 11 2533–2541. the original title, "Practical machine learning", was changed ... The term "data mining" was [added] primarily for marketing reasons.
- [10] J Mena, Jesús (2011). *Machine Learning Forensics for Law Enforcement, Security, and Intelligence*. Boca Raton, FL CRC Press (Taylor & Francis Group). ISBN 978-1-4398-6069-4.
- [11] Piatetsky-Shapiro, Gregory; Parker, Gary (2011). "Lesson Data Mining, and Knowledge Discovery An Introduction". *Introduction to Data Mining*. KD Nuggets. Retrieved 30 August 2012.
- [12] Kantardzic, Mehmed (2003). *Data Mining Concepts, Models, Methods, and Algorithms*. John Wiley & Sons. ISBN 0-471-22852-4. OCLC 50055336.
- [13] "Microsoft Academic Search Top conferences in data mining". Microsoft Academic Search.
- [14] "Google Scholar Top publications - Data Mining & Analysis". Google Scholar.
- [15] *Proceedings, International Conferences on Knowledge Discovery and Data Mining*, ACM, New York.
- [16] *SIGKDD Explorations*, ACM, New York.