A Survey On Smart System For Student Placement Prediction

Madhura B Shetty, L.Pratibha, Lavanya.K , Chaitra B S, Sachin D N
Department Of Computer Science& Engineering
Vidyavardhaka college of Engineering, Mysuru

Abstract: ***The biggest defies that higher education institutions look currently is to increase students placement percentage. Educational institutes are in search of new technology that can increase the placement status of the institution. One efficient ways to improve the objective is to come up with the new acquaintance connected to the learning methods and objects. Using machine learning methods the information can be inferred after effective and chronological records which will be in the educational institute database. For the student implementation the datasets is of the information of the students from the past year. For rule identification the data is used for training model and for classification it is the testing model. This gives a reference system that foresees the students to get placed to any of these category i.e. Core Company, Dream Company, Mass Recruiters and Not Interested in Placements. This system supports the placement cell within an educational institute to categorize the potential of the students and consideration to increase students methodological and social skills. Further, the prefinal and final years students of the B.E course can access this system to identify their position in placement individually. Through this they can come up with more hardwork for getting placed in to the firms that belong to greater hierarchies.***

Keywords: Machine leaning, Naïve Bayes, Data Mining, Clustering, Placement.

INTRODUCTION

The premier aim of students joining proficient courses in universities is to get placed to a lucrative job in top companies. Qualified education can be of entirely methodological or it can be decision-making. B.E provides methodical training to students in numerous grounds such as Computer Science, Electronics and Communication, Civil, Mechanical, Electrical and electronics Engineering. etc. This B.E degree is intended to make students professionals in hypothetical and practical knowledge in different fields of engineering. Since the algorithm experts in state of the art academic and hands-on acquaintance in various fields of engineering the competence is to analyze the prediction of placement position will help the B. E students most likely to attain and to make suitable advancement in their career field. It will also help the faculties

Through the outcomes it was found ID3 algorithm is suitable for predicting the placement of students in which ID3 gives around 85.33% prediction compared to C4.5 and CHAID. [2]

and placement cell in an organization to deliver appropriate supervision to the perfection of students during their course. The increase in percentage of placement is the primary key in constructing the status of an educational organisation. Therefore such a system has a considerable position in the educational system of any developed education institute.

LITREATURE SURVEY

In this paper the method used is sum of difference which is used to predict the student placement analysis. Gender, Category. Academic gap , 10^{th} Grade, 12^{th} Grade , Number of arrear, Grade in BTech exam ,communication skills in English are taken as trait. The data is gathered from 50 students with the listed attributes. For each attribute chosen

, a reference point is given in between 0.0 to 1.0 on its precedence. Next another reference point is chosen for each trait in the range of value of given trait and detract it with each value correspondingly. The work is dealt with the addition of each subtracted value. Based on the value, a end point value is initiate where placement value can be reformed. The overall result was, when the sum of difference value rise till 4; it specifies no value of placement. When the value goes high than 4, it specifies yes value of placement. [1]

This paper is implemented using decision tree algorithm. Here they have used ID3,C4.5 and CHAID algorithm. Rapid Miner technology is used. The attributes selected are roll number, HSG, UG, Board and communication In this case, they have collected 1342 students details as a dataset from placement.Supervised learning is used in this paper where each node is divided using the subset of examples at the node which is selected randomly. If sem is good, Internal is good, Medium is English then performance will be good. Similarly if hindi mid sem is poor and internal is poor then the performance will be average deprived. If the medium is average then mid sem is deprived and internal is reduced .Then they analyze the performance based on the accuracy.

In this paper the data mining technique was adopted. They have opted for clustering technique which is done through

K-Means algorithm. The study is achieved using Rapid Miner studio. Methodology implemented in this paper is that statistics for a segment of student in computer science of class 2012-2016.Attribute such as Xth, XII and graduation marks were considered. In addition to this projects, internships were also considered as relevant. The number of students in cluster 0, cluster 1, cluster 2 were 12, 37 and 50.when the analysis were made on all the attributes of the all three clusters0, 1 and 2 the performance of students based on collective study was 5.36, 6.00, and 6 2 respectively. It was inferred the number had a better performance while merely 40 poor performance in cluster 1. [3]

This paper lead a study for the characteristics affecting the students academic performance. They used questionnaires to elicit info from students prominence aspects like income factor, parents' informative background, consistency of teachers, subject interest of teachers and students cocurricular activities. The method used is Pearson Correlation Coefficient to highlight the significant features. The outcome was that mother's education and family income played a vital key role in academic performance of students. [4]

This paper describes about the student placement prediction using data mining classification rules and regression models. Basically Fuzzy algorithm is used for classification. Here they consider a dataset of passed out students which contains the information of four years aggregate and the company they got placed to. Linguistic variables considered here are 'Highest', 'High', 'Moderate', 'Low' and 'Lowest'. A matrix of current students and placed students attribute values is considered. A set of rules which are in the form IF-THEN-ELSE structures are used. Fuzzy set operations perform estimation of these rules. The operations MAX and MIN are OR and AND respectively. Finally the output is the Fuzzy value. [5]

In this survey they have used distributed data mining technique for extracting the information of students in the form of diagrams and arrangements from collected data. The results obtained from this type of technique will be more accurate which can be used for decision making purposes. The performance of the system increases by using this technique. Random tree algorithm which comes under is poor, backlog is yes then the performance will be poor.

If the sem is meager and backlog has yes status then the performance will be poor. The over all result is that random tree decision algorithm gives more accuracy when compared to other decision tree algorithms and reduce the time required for classification, hence it is efficient. [6]

This paper led a study on statistics of student that ensure data on their educational records and expected a classification to discover an resourceful technique to expect placements of student. The overall result was that Naïve Bayes classifier is the best technique for use in placements in contrast with algorithms such as Multilayer Perceptron and J48. [7]

The paper for the extrapolation of placement in campus using the two Data Mining algorithms and they are Fuzzy Logic and K nearest neighbor" (KNN). Firstly the dataset of students will be calm from the respective educational institutions. The dataset comprises PU marks, UG aggregate and personality development skills of the student like seminar skills, practical skills, etc. Here the data of 900 students is considered where 600 dataset is used for training set to build the model and remaining is the testing data for authorizing the model. At first a classifier is built with the training data and next is the classification stage. Using both Fuzzy Logic and K nearest neighbor, the output is given on the basis of the accuracy of the two models. [8]

LITERATURE REVIEW

AUTHOR NAME	METHOD	ANALYSIS	ACCURACY
1.Ramanathan, L., P. Swarnalathat G.D. Gopal	Sum of Difference	when amount of variance rate rise till 4 it says no rate of placement. When the rate is high than 4, it specifies yes value of placement.	60%
2. T.Jeevalatha, N,Ananthi and D.SaravanKumar	ID3,C4.5 & CHAID	Among the 3 methods ID3 is more appropriate for placement prediction.	85.33%
3. Ishawank singh,A Sai Sabitha and Abhay	K-Means	Comparing 3 clusters cluster 1 is	60%

		accurate.	
4. .Hijazi, S.T. and R. S. M. M. Naqvi	PCC	Attributes are not appropriate	< 40%
5.Raina Sangha,Akshay satras,Lisha Swamy&Gopal Deshmukh	Fuzzy	Considered attributes are less for prediction.	75%
6.Krina Parmar,DineshKumar Priyanka sharma	RDT	Time complexity is less.	70%
7. Pal, A.K. and S. Pal	Naïve Bayes	Results were more accurate.	92%
8. Mangasuli Sheetal Savita Bakare	KNN & Fuzzy	Attributes extends to personality development skills.	87%

COMPARITIVE ANALYSIS

As of now there are many methods found to predict the student placements. We have come across the methods like 'sum of difference', 'KNN', 'K-Means', Fuzzy, Clustering techniques, Pearson Correlation Coefficient and Naive Bayes. In sum of difference, when sum of difference rate rise till 4 it says no value of placement. When the rate is high than 4, it specifies yes value of placement. Hence the results may not be that appropriate. In ID3, C4.5, CHAID methods ID3 is more appropriate for placement prediction. The results may be accurate upto 85.33% .In K-Means cluster 0,cluster 1,cluster2 comparing these three clusters, cluster 1 had appropriate prediction .In Pearson correlation coefficient attributes chosen were not appropriate. Hence accuracy is less than 40%.In Fuzzy method attributes considered for the prediction were less.In Random Decision Tree time complexity is less hence prediction accuracy is faster. In KNN and Fuzzy method attributes extends to personality development skills and accuracy is upto 87%. Naïve Bayes is based on group of classifiers method .Hence efficiency and results are more appropriate. After the survey of all the papers for student placement prediction it is found that Navies Bayes method has more appropriate result with more accuracy.

[7]. Pal, A.K. and S. Pal (2013)Analysis and Mining of Educational Data for Predicting the Performance of Students. (IJECCE) International Journal of ElectronicsCommunication and Computer Engineering, Vol. 4, Issue 5, pp. 1560-1565, ISSN: 2278-4209, 2013.

CONCLUSION

At current study in informative data mining it creates the awareness in the study communal. In most of the cases, 10th, 12th and CGPA are considered as significant attributes for estimation of result. The prediction, classification and clustering technique is often used procedure. Most of the researchers used Decision Tree, sum of difference, Naive Bayes and Fuzzy C-Means. At last, we can accomplish that the meta-analysis on foreseeing placement performance inspired us to do future research for our individual informative atmosphere. It would help to increase our education organisation towards patterned the placement of the students.

REFERENCES

[1]. Sentkil Kumar Thangavel, P. Divya Bkaratki, Abijitk Sankar. "Student placement analyzer: A recommendation system using machine learning", 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017

[2].Ramanathan, L., P. Swarnalathat and G.D. Gopal (2014)Mining Educational Data for Students' Placement Prediction using Sum of Difference Method. International Journal of Computer Applications 99(18): 36-39

[3].T.Jeevalatha,N,AnanthiandD.SaravanKumar(2014) Performance analysis of Undergraduate studentsPlacement Selection Using Decision Tree Algorithm vol108-No 15,DECEMBER 2014

[4]. Singh, Ishwank, A Sai Sabitha, and Abhay Bansal. "Student performance analysis using clustering algorithm", 2016 6th International Conference - Cloud System and Big Data Engineering (Confluence), 2016.

[5] Raina Sangha,Akshay satras,Lisha Swamy&Gopal Deshmukh Students Placement Eligibality prediction using Fuzzy approach,International journel of Engineering and techniques-vol Dec-2016.

[6] Parmar, Krina, Dineshkumar Vaghela, and Priyanka Sharma. "Performance prediction of students using distributed Data mining", 2015 International Conference on Innovations in Information Embedded and Communication Systems (ICIIECS), 2015.

[8] Mangasuli Sheetal Savita Bakare Prediction of Campus Placement Using Data Mining Algorithm-Fuzzy logic and K nearest neighbour International Journal of Advanced Research in Computer and Communication Engineering Vol. 5,
Issue 6, June 2016

[9].Rokach, L and O. Maimon (2008)Data mining with decision trees: theory and applications.World Scientific Pub Co Inc. ISBN 978-98127717711

[10]. S. Rajeswari, R. Lawrance. "Classification model to predict the learners' academic performance using big data", 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), 2016

[11].Abijith Sankar, P. Divya Bharathi, M. Midhun, K. Vijay, and T. Senthil Kumar, "A Conjectural Study on Machine Learning Algorithms," In Advances in Intelligent Systems and Computing published – Springer - from the proceedings of International Conference ICSCS 2015 2016, Vol.397, pp. 105-116.

[12].Elayidom, S., S. M. IdikkulaandJ.Alexander(2011)A Generalized Data mining Framework for Placement Chance Prediction Problems.International Journal of Computer Applications(0975– 8887) Volume 31– No.3.