RESEARCH ARTICLE   OPEN ACCESS

Student's Placement Eligibility Prediction using Fuzzy Approach

Ravina Sangha[1], Akshay Satras[2], Lisha Swamy[3], Gopal Deshmukh[4]
[1,2,3,4]*(Department of Computer Engineering, SPPU, Pune)*

Abstract:
      The main objective of this paper is to develop a basic prototype model which can determine and extract unknown knowledge (patterns, concepts and relations) related with multiple factors from past database records of specific students. Data mining is science and engineering study of extracting previously undiscovered patterns from a huge set of data. Data mining techniques are helpful for decision making as well as for discovering patterns of data. In this paper students eligibility prediction system using Rule based classification is proposed to predict the eligibility of students based on their details with high prediction accuracy. In Educational Institutes, a tremendous amount of data is generated. This paper outlines the idea of predicting a particular student's placement eligibility by performing operations on the data stored. In this paper an efficient algorithm with the technique Fuzzy for prediction is proposed.

Keywords — ***Classification, Prediction, Fuzzy***.

I.   INTRODUCTION

   An educational system has huge quantity of information regarding students, staff, corporations etc. currently a way to analyze this information to create higher results from information we tend to use totally different techniques. That's the employment of information of knowledge mining technique which implies extracting helpful data from the galore quantity of information. By applying varied techniques we will predict the results.

   There are varied techniques for predicting like classification, regression and attribute importance. There are varied model constitute these classes. As in classification we've boosted classification tree, support vector machine, naive Thomas Bayes, C4.5 etc. these all are used for classification. In same manner we've totally different models for regression that are model tree, neural network, strong linear model, random forest generator, Wang and plant scientist fuzzy rules. By creating ensemble of these models prediction may be performed.

   The prediction of eligibility of scholars in varied corporations is additional necessary and this could be accomplished by this model. This can facilitate everybody as staring from students they'll brace oneself for corporations before, they'll improve their CGPA to

eligible for corporations, predict their future career and etc. currently academics can even facilitate students in varied fields and improve their skills. And this reciprocally can facilitate in rising the name of an establishment in IT education. The present study concentrates on the engineer students for his or her billet in pre-final year and placement in final year graduates moreover as post graduate students. By applying data processing techniques the results are shown.

II.   RELATED WORK

   Stamos T. Karamouzis and Andreas Vrettos [1] state that declining student commencement fees is a significant and developing problem in better education. College students are dropping out from schools for an expansion of motives and college directors are scrambling to growth commencement costs. Predicting pupil commencement is of incredible cost to colleges and a full-size capacity utility for focused intervention. Thinking about the promising conduct of synthetic Neural Networks (ANNs) as classifiers led us into the improvement, training, and checking out of an ANN for predicting pupil graduation consequences. The network turned into evolved as a three-layered notion and was trained using the again-propagation standards. For

training and testing various experiments had been performed. In these experiments, a pattern of 1,407 profiles of college students became used. The pattern represented college students at War Bonnet College and it turned into divided into two units. The primary set of 1, a hundred profiles changed into used for training and the closing 307 profiles had been used for trying out. The common predictability fee for the schooling and take a look at units had been seventy seven% and sixty eight%, respectively.

Data mining provides the ideas and techniques in processing gathered knowledge or data, which can be employed in numerous applications. Specifically, it explains data processing and also the tools employed in discovering data from the collected data. [2] Focusses on the data discovery from knowledge (KDD). It focuses on the feasibility, usefulness, effectiveness, and quantifiability of techniques of enormous knowledge sets. After describing data processing, this edition explains the ways of knowing, preprocessing, processing, and repositioning knowledge. It then presents data concerning knowledge warehouses, online analytical process (OLAP), and knowledge cube technology. Then, the ways concerned in mining frequent patterns, associations, and correlations for giant knowledge sets are delineated.

Tongshan Chang Jiang [3] introduces a case study victimization data processing techniques to help instruction institutions in achieving enrolment goals. The introduction includes the final life cycle of an information mining project: business understanding, knowledge understanding, knowledge preparation, modelling, assessment, and preparation. The model match statistics show that data processing techniques, neural network, call tree, logistical regression, and ensemble models, are winning during this project. Ensemble and neural network are higher. The study concludes that data processing is a good technology for school achievement and conjointly for institutional research and analysis. Data processing, outlined as "the method of sampling, exploring, modifying, modelling, and assessing giant amounts of information to uncover antecedently unknown patterns", has been used wide in numerous areas like science, engineering, business, banking, and even combating act of terrorism for an extended time. Its success and effectiveness in discovering unjust data from giant sets of information has been well established.

In [4] Data mining is broadly used for selection making in instructional device. One of the maximum drastically used supervised getting to know strategies i.e. decision tree classifier for records exploration based

totally on divide & conquers method. This paper discusses use of decision timber in instructional facts mining. Choice tree algorithms are implemented on company's preceding 12 months statistics to generate the model and this model may be used to be expecting the students' eligibility in various companies. This could assist students to identify the category of enterprise in which they're eligible and prepare as a consequence in an efficient way.

Data mining is this sort of promising era whose worth becomes obtrusive when it is able to be applied to a domain wherein a common guy is benefited. [5] is a try to assist the prospective students to make smart profession decisions using technologies like data mining the usage of selection trees, Naïve Bayes and artificial neural networks. A student enters his front Rank, Gender (M/F), region (rural/urban) and Reservation category. Based on the entered facts the community or the choice tree will return which department of examine is incredible, properly, average or bad for him/her. Additionally in this paper we compare the overall performance of the fashions on the equal information and advice a generalized records mining framework for issues of similar nature.[6] Explores the socio-demographic variables (age, gender, ethnicity, schooling, paintings status, and disability) and has a look at environment (course programmed and course block), that may affect persistence or dropout of college students at the Open Polytechnic of New Zealand. We look at to what quantity those factors, i.e. enrolment records assist us in pre-identifying successful and unsuccessful college students.

Many academic institutions want their college students to be placed in reputed agencies and therefore they recognition on the educational overall performance of their college students. If they might able to recognize approximately the scholar's instructional performance, then they can recognition greater on those students whose educational performance is terrible. They can take unique care on the ones college students whose academic performance are negative and make them prepare for the location. In [7], we present a look at on the academic performance of the students of a reputed university with the aid of applying k-method set of rules of Data Mining which allows them to understand approximately the performance of their students.

Nadeem Talib, Sujit S. Sansgiry [8] within the gift observe, elements consisting of educational competence, take a look at competence, time control, strategic studying, and test tension were studied as determinants of instructional overall performance, i.e., Grade point common. Similarly, these aspect may also help in

identifying high in addition to low educational achievers. A sample of 199 undergraduate and graduate college students from Rawalpindi and Islamabad acquired a survey with the changed model of the study control and educational effects test (Topman, Kleijn, Ploeg, & Masse!, 1992) and the test tension stock (Sarason, 1980). The effects indicated that instructional competence, take a look at competence, time management, and take a look at tension were notably associated with student's educational performance. Outcomes additionally confirmed that check competence, academic competence, and check anxiety being the primary discriminators amongst low and excessive GPA achievers. developing techniques that help college students deal with the rigors of educational existence, recognize how to take a look at successfully for assessments, and assisting them reduce their level of hysteria related to taking a test could help improve their destiny overall performance, these advantages might be visible specifically for college students who have a low GPA.

Studying the engineering pupil educational performance isn't always an easy challenge for the network of excessive studying. The performance of pupil for the duration of their first year in university is a turning factor of their educational direction and generally encroaches on their final percent in a decisive manner. The scholar's evaluation elements like magnificence quiz, internals, and overall performance in the lab could be studied. Analyzing [9] will assist the academics to reduce drop out ratio to a large level and enhance their performance within the final examination. Information performs an crucial position in analyzing and comparing the overall performance in university to be able to make suitable instructional choices. Educational selections will end result adjustments in their performance which want to be assessed periodically over span of time. The overall performance parameters chosen can be viewed at character scholar, elegance, and department and college degrees. Most important clustering strategies are used to extract meaningful information and to expand the enormous relationship a few of the variables stored in massive statistics sets. In this paper, we gift a process primarily based on choice tree of statistics mining techniques and ok-method partitioning of clustering methods which allows to beautify the first-class of educational system by means of reading and improving student's performance.

Siu-man Raymond Ting [10] state that the primary year in college adjustment and turmoil for many late youth. A few experience difficulties sufficient to cause

them to drop out. Academic performance and retention of college students has been studied drastically and theoretical model broaden to explain various factor affecting college students adjustment and educational overall performance. At gift, the Scholastic aptitude test (SAT) is a typically used standardized take a look at for under-graduate admission inside the America for improving overall performance of college students in educational.

III.   PROPOSED WORK

Let S denote the system under consideration,
S= {I, O, F, Su, Fa}
I = Defines the input given to the system
I = {U, Q, D}
U = any random User
U = {u1, u2….un}
Q = User will enter the query
Q = {s1, s2, s3…sn}
D = The Dataset
F = Functions used for classification {Fuzzy Logic}
Su = Success if the predicted result is accurate
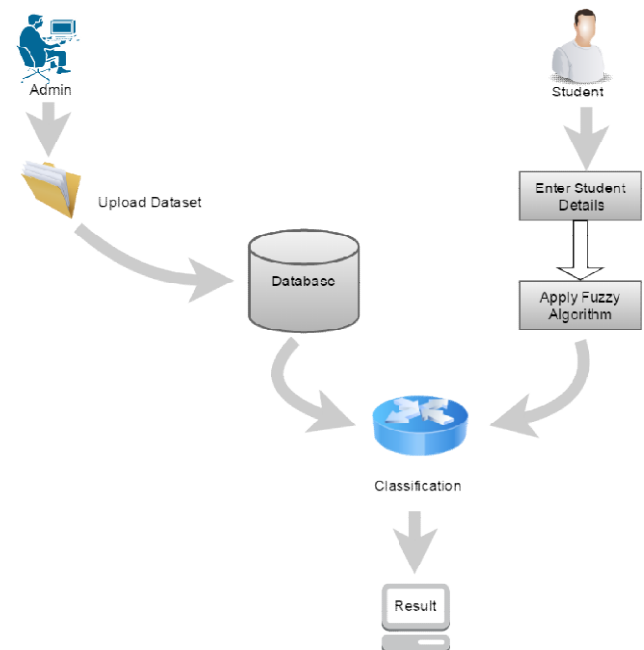Fa =Failure, if the result is inaccurate

Fig. 1 System Architecture

A.   Data Preparation

The dataset will be already passed out student, that dataset contains four year aggregate and in which company that student has been placed.

B.   Data Selection and Transformation

In second step the useful information is being selected according to requirements. There are original fields which are taken as original variables and from that derived variables are generated. The file can be uploaded and the results can be generated. File data can be changed and uploaded any time.

IV. FUZZY ALGORITHM

Step 1: Define linguistic variables and terms
Linguistic variables are input and output variables that are in terms of words. For student eligibility prediction highest, high, moderate, low, lowest, etc., are linguistic terms.
Eligibility (e) = {highest, high, moderate, low, lowest}
Every member of this set is a linguistic term and it can cover some portion of overall values.

Step 2: We need to construct membership functions for the linguistic variables.

Step 3: Construct the knowledge base rules
Create a matrix of current student's attribute values versus target attribute values of previously placed students which correspond to the value that is expected.

Input/Target	Highest	High	Moderate	Low	Lowest
Highest	No Change	No Change	No Change	No Change	No Change
High	No Change	No Change	No Change	No Change	No Change
Moderate	Slight Improvement	Slight Improvement	No Change	No Change	No Change
Low	Strong Improvement	Strong Improvement	Strong Improvement	No Change	No Change
Lowest	Strong Improvement	Strong Improvement	Strong Improvement	Strong Improvement	No Change

Table 1 : Matrix Representation

Construct a set of rules into the knowledge base which are in the form of IF-THEN-ELSE structures.

Sr No.	Condition	Action
1	IF eligibility=(Highest OR High) AND target=High THEN	No Change
2	IF eligibility=(Lowest OR low) AND target=High THEN	Strong Improvement
3	IF (eligibility=Moderate) AND (target=High) THEN	Slight Improvement

Table 2:Condition Action Table

Step 4: Obtain Fuzzy Value
Fuzzy set operations perform evaluation of rules. The operations used for MAX and MIN are OR and AND respectively. Finally combine all results of evaluation to generate final result. This result is a fuzzy value.

Step 5: Perform Defuzzification
Defuzzification is then performed according to membership function for output variable.

V. ADVANTAGES

The main advantage of Fuzzy logic is that it gives you a variety of possibilities and not just one fixed value. It gives you the possible values for any predicted result between 0 and 1.At the same time it is robust as well, that means it can handle erroneous input in a very efficient manner. Since the algorithms can be described with a little amount of data, it does not require much space in the memory. Fuzzy logic can give effective results even with limited knowledge i.e. similar to human beings.

Fuzzy can also deal with non-linearity and models your reasoning. The most significant advantage of this approach is that it can deal with linguistic variables. Linguistic variables are variables expressed in terms of words. Then for each member of the set, a membership function is defined. In short, Fuzzy logic is a solution to complex problems in many fields of life, including medicine, as it resembles human reasoning and decision making.

DATASET
The dataset on which this algorithm will be tested will be that of MESCOE Pune. Training data will also be of the same source which will be provided by the Training and Placement Cell administrator. Thus, actual and predicted value will be verified.

VI. FUTURE SCOPE
By using Fuzzy algorithm with K means clustering algorithms quicker and faster system can be developed. This is possible because K means is capable of forming clusters of students with similar properties and thus Fuzzy can assign probability values to the entire cluster instead of assigning probabilities individually to each and every student.

VII. CONCLUSION
Thus, this system will help the student in getting placed in a company which he is capable for and in turn

will help the institutes to improve their student placement record rate which will certainly increase the number of applicants who desire to get educated in such a college.

REFERENCES

1. *S.T.Karamouzis S.T., and Vrettos, A., 2008. An Artificial Neural Network for Predicting Student Graduation Outcomes, Proceedings of the World Congress on Engineering and Computer Science, pp. 991-994..*

2. *J. Han and M. Kamber, "Data Mining: Concepts and Techniques," Morgan Kaufmann, 2000..*

3. *T .Chang, Ed. D, "Data Mining: A Magic Technology for College Recruitment´s," Paper of Overseas Chinese Association for Institutional Research (www.ocair.org), 2008*

4. *M. Gera and S. Goel, "A model for predicting the eligibility for placement of students using data mining technique," International Conference on Computing, Communication & Automation, Noida, 2015, pp. 114-117.*

5. *S. Elayidom, S. Mary Idikkula & J.Alexander, "A Generalized Data mining Framework for Placement Chance Prediction Problems," International Journal of Computer Applications, (0975– 8887) vol.31, no. 3, October 2011.*

6. *Z.J. Kovacic, "Early prediction of student success: Mining student enrollment data," Proceedings of Informing Science & IT Education Conference, 2010.*

7. *G.N. Rao, M. Ramachandra, "A Study on the Academic Performance of the Students by Applying K-Means Algorithm," International Journal of Emerging Technologies in Computational and Applied Sciences, vol. 7, no. 4, pp. 365-366, February 2014.*

8. *Talib, Nadeem, and Snjit S. Sansgiry. "Determinants of academic performance of university students." Pakistan Journal of Psychological Research 27.2 (2012): 265.*

9. *Analyzing the Student's Academic Performance by using Clustering Methods in Data Mining Sreedevi Kadiyala, Chandra Srinivas Potluri International Journal of Scientific & Engineering Research, Volume 5, Issue 6, June-2014.*

10. *Ting, S. M. R., and R. Man. "Predicting academic success of first-year engineering students from standardized test scores and psychosocial variables." International Journal of Engineering Education 17.1 (2001): 75-80.*