Placement Recommender and Evaluator

Darshan Yadav[1], Omkar Shinde[2], Anup Singh[3], Asst Prof. Asmita Deshmukh[4]

[1,2,3]Student, K.C College of Engineering & Management Studies & Research, Thane- 400603, Mumbai
[4]Assistant Professor, Dept. of Computer Engineering, K.C College of Engineering & Management Studies & Research , Maharashtra, India

---

Abstract - *A college campus recruitment system that consists of a student login and an admin login. The project is beneficial for college students, various companies visiting the campus for recruitment and even the college placement officer. The software system allows the students to create their profiles and upload all their details including their marks onto the system. The admin can check each student details and can remove faulty accounts. The system also consists of a company records where various companies visiting the college can view a list of students in that college and also their respective resumes. The software system allows students to view a list of companies who have posted for vacancy(which is updated by the TPO). The admin has overall rights over the system and can moderate and delete any details not pertaining to college placement rules. The system handles student as well as company data and efficiently displays all this data to respective sides. The students each have been given information about domains. The domain-student compatibility is checked using thorough questions, aptitude tests, etc. Each student can choose up to five domains of their choice after which the PRE will check and select the best possible domain. The records of each alumina are stored in case to obtain more detailed information about the respective company.*

Key Words: **Recruitment**, **Compatibility**, **Vacancy**, **Efficiency, Domains, Aptitude.**

1. INTRODUCTION

The performance in education sector in India is a turning point in the lives of all students. As this academic performance is influenced by many factors, it is essential to develop predictive data mining model for students' performance so as to identify the slow learners and study the influence of the dominant factors on their academic performance.

Campus placement of a student plays very important role in a college. Campus placement is a process where companies meet colleges and identify students which are talented and qualified, before they complete their graduation

The educational sector in IT includes the student records namely aptitude skills, certification courses, technical abilities in various languages or web development and academic performance.

Students studying in final or third year of an Engineering college start feeling the pressure of the placement season with so much of placements activities happening around them. They feel the need to know where they stand and how they can improve their chances of getting job. The Placement Office plays a important role in this. The students are given vital information on how to prepare themselves for the placement season by the TPO.

It may be an important consideration to analyze various trends since all the systems are now computer based information system so data availability, modification and updating are a common process now. Student achievement is highly influenced by past evaluations which involves various relevant features (e.g. attendance in lectures/practical, participation in various intercollegiate college events, test scores, etc.). As a direct outcome of this project, more efficient student prediction tools can be developed, improving the quality of education and enhancing school resource management.

Student prediction system is most important approach as it classifies large set of student data is very difficult so many organizations suffer from this problem and this would be serious concern.

If we are going to classify large set of data set of student on excel sheet it will take lot of time and if we use any programming language, it is also very difficult to code for many conditions and classify big data set. This is the serious issue which frequently occurs during classification of student big data set or predicting them for desire purpose.

The eligibility criteria of students in various companies is more important and this can be realize by this model. This will help everyone as beginning from students they will prepare for companies in advance. the objective of TPO management system is send campus interview notification to those candidate who are eligible for that.

For this we will consider the academic history of the student like percentage as well as their skill set like, programming skills, communication skills, analytical skills and team work, which are tested by the hiring companies during the recruitment process.

Though having many systems/platforms where you can judge where you stand, but not under a single shed where a user can get reviewed. Having such a system might get users more directed to their goal with knowing the current corporate requirements.

In colleges when the placements arrive they intimate the students by posting the information on notice board so if any of the students are not attending the college. He cannot know the information about the placements. so this is the drawback of the existing system.

- Though having many systems/platforms where you can judge where you stand, but not under a single shed where a user can get reviewed.
- Having such a system might get users more directed to their goal with knowing the current corporate requirements.
- Institution can take the generated report as a feedback, and can be helpful in profiling the students qualities, what are his/her weak sections or where he needs improvement or by considering this as his future coming domain won't be a wrong choice.
- Students can evaluate themselves in any of their fields of choice. Included AI in the system will help them take a wise decision.

1.1 Data Mining

- Data Mining is the process of extracting useful information from large sets of data.
- Data mining enablesthe users to have insights into the data and make useful decisions out of the knowledge mined from databases.
- The purpose of higher education organizations is to offer superior opportunities to its students. As with data mining, now-a-days Education Data Mining (EDM) also is considered as a powerful tool in the field of education.
- It portrays an effective method for mining the student's performance based on various parameters to predict and analyze whether a student (he/she) will be recruited or not in the campus placement.
- Predictions are made using the machine learning algorithms J48, Naïve Bayes, Random Forest, and Random Tree in weka tool and Multiple Linear
- Regression, binomial logistic regression, Recursive Partitioning and Regression Tree (rpart), conditional inference tree (ctree) and Neural Network (nnet) algorithms in R studio.
- The results obtained from each approaches are then compared with respect to their performance and accuracy levels by graphical analysis.
- Based on the result, higher education organizations can offer superior training to its students.

2. PROPOSED SYSTEM

- Our proposed system, can help institution analyze the technical skill sets of students and can be helpful in predicting the path of students and it can earn companies to analyze the quality of students and thereby eliminating unnecessary rounds for incapable candidates.
- At the beginning of the second year when students first enter the course they are interested, detailed information is provided to the students on the course they chose and what future jobs they can choose.
- Here based on students level of study he/she can get him/herself evaluated by undergoing series of evaluation tests. At the very start each student has to have to give a simple aptitude test to measure their level of thinking (because it is mandatory in almost every recruitment drives).

- Each evaluation test has domain specific questions prepared thoroughly for enhancing the domain specific skills.
- Each student will be provided with proper guidance regarding the competitions(with their credit for future) and all certifications(with their credit for future) from the second year of their college better informational guidance.
- One of the biggest challenges the higher education institution faces today is predicting the paths of students.
- In PRE System, test reports can be used by institution to evaluate the students performance in each test of his/her level for profiling purpose especially in the early years of their study.
- Using *URJob* module candidate can get self evaluated based on his/her skill set and get to know where he/she stands in todays competitive world.
- As a feedback, student will get *Tutorials* notifications based on his/her performance in test and his/her scope of improvement areas.
- The GUI is done using Django which is a python web development framework.
- Machine Learning (for generating a pattern in the sequence of earlier inputs).
- Data Mining.(Data Mining is the process of extracting useful information from large sets of data. Data mining enables the users to have insights into the data and make useful decisions out of the knowledge mined from databases. The purpose of higher education organizations is to offer superior opportunities to its students. As with data mining, now-a-days Education Data Mining (EDM) also is considered as a powerful tool in the field of education.)
- AI (for making system self-reliant to predict a particular domain for a student).

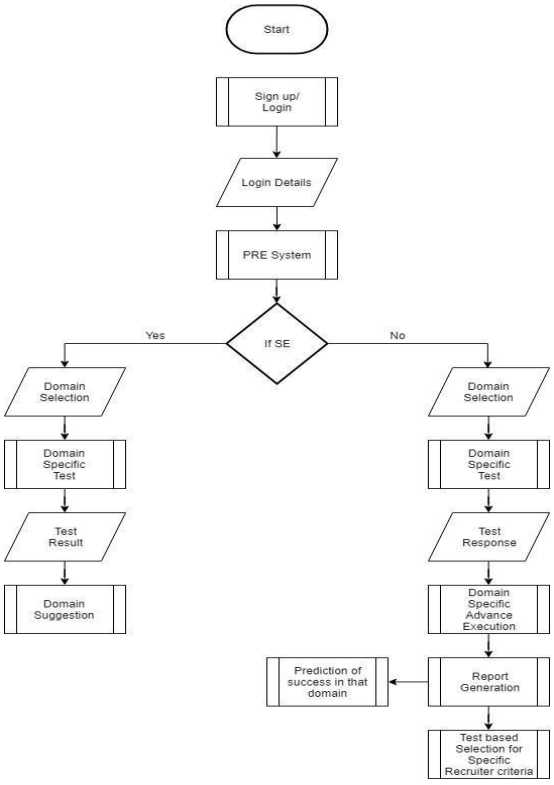

Fig-1: System Flow Chart

3. C4.5 Classification Algorithm

Handling both continuous and discrete attributes - In order to handle continuous attributes, C4.5 creates a threshold and then splits the list into those whose attribute value is above the threshold and those that are less than or equal to it.

Handling training data with missing attribute values - C4.5 allows attribute values to be marked as ? for missing. Missing attribute

values are simply not used in gain and entropy calculations.

Handling attributes with differing costs.

Pruning trees after creation - C4.5 goes back through the tree once it's been created and attempts to remove branches that do not help by replacing them with leaf nodes.

```
J48 pruned tree
------------------

Gender = Male
|   SSC <= 79: TCS (2.0)
|   SSC > 79: Wipro (2.0)
Gender = Female: Infosys (2.0)

Number of Leaves  :     3

Size of the tree :      5


Time taken to build model: 0 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0 seconds

=== Summary ===

Correctly Classified Instances           6               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances                6
```

Fig-2: Classifier Output

```
=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     TCS
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Wipro
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Infosys
Weighted Avg.  1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

 a b c   <-- classified as
 2 0 0 | a = TCS
 0 2 0 | b = Wipro
 0 0 2 | c = Infosys
```

Fig-3: Algorithm Output

4. RESULTS

The Placement Recommender and Evaluator system applies data mining techniques using various decision tree and Naïve Bayes classifiers. Decision treesare easy to understand models as they we can generate separate module for each conclusion.

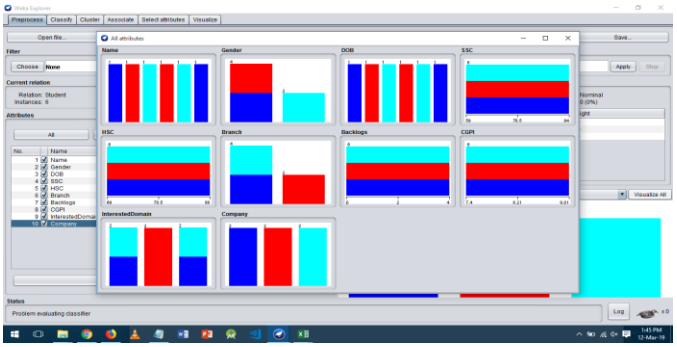

Fig-4 : Classification using WEKA Tool

Also as the UI is completely user friendly one can easily have great user experience while evaluating himself.

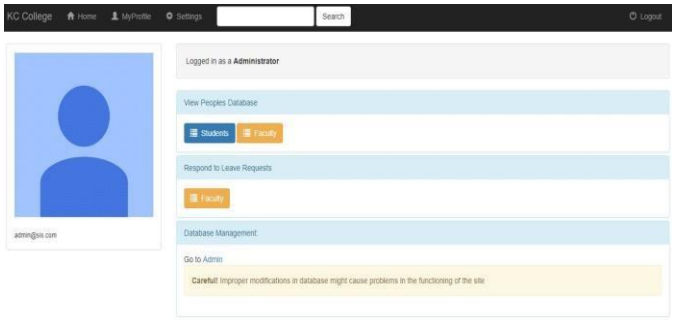

Fig-5 : User Interface Design

5. CONCLUSIONS

Our proposed system, can help institution analyze the technical skill sets of students and can be helpful in predicting the path of students and it can earn companies to analyze the quality of students and thereby eliminating unnecessary rounds for incapable candidates.

To conclude, we will be predicting the placement results using C4.5 algorithm and Naïve Bayes Classifier. In educational field, C4.5 gives much better prediction than any other classification algorithms. We have observed that as the number of data increases the prediction accuracy using C4.5 also increases. This system can be further used in various college's placement cells to help enhance the placement procedure.

REFERENCES

[1] Bhullar, Manpreet Singh, AmritpalKaur, "Use of Data Mining in education sector."Proceedings of the World Congress on Engineering and Computer Science.Vol. 1. 2012.

[2] Ajay Kumar Pal, Saurabh Pal "Classification Model of Prediction for Placement of Students" I.J.Modern Education and Computer Science,2013,11,49-56.

[3] http://gerardnico.com/wiki/data_mining/naive_ba yes .

[4] S. Nagaparameshwara Chary, Dr.B.Rama "Analysis of Classification Technique Algorithms in Data mining".

[5] Ajay Kumar Pal, Saurabh Pal "Classification Model of Prediction for Placement of Students" I.J.Modern Education and Computer Science, 2013, 11, 49-56.

[6] Sujith Jayaprakash, Balamurugan E., VibinChandar "Predicting Students Academic Performance using Naive Bayes Algorithm".

[7] http://gerardnico.com/wiki/data_mining/naive_ba yes.

[8] Witten, I.H. & Frank E., Data Mining– Practical Machine Learning Tools and Techniques, Second edition, Morgan Kaufmann, San Francisco, 2000.

[9] S. Nagaparameshwara Chary, Dr. B.Rama "Analysis of Classification Technique Algorithms in Data mining-A Review".

[10] Cortez, Paulo, and Alice Maria Gonçalves Silva. "Using data mining to predict secondary school student performance." (2008).

[11] Zaíane, Osmar R. "Building a recommender agent for e-learning systems." Computers in education, 2002. Proceedings International conference on. IEEE, 2002.

[12] J. Han and M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2000.