Student Placement Prediction Using Weka

1 Jaimi Santhosh 2 Navyamol K.T
1 PG Scholar of Amal Jyothi College of Engineering, Kanjirappally, Kerala
2 Assistant Professor of Amal Jyothi College of Engineering, Kanjirappally, Kerala
1 jaimisanthosh@mca.ajce.in 2 navyamolkt@amaljyothi.ac.in

Abstract: Nowadays, placement is one of the important activities in educational institutions. The key role that the educational institution is to prepare and guide students well through the placement process by providing them better training for getting placed in reputed companies. Best placement service of an institution will also attract people to take admissions. A placement prediction system helps to know the possibility of a student to be placed. Machine learning techniques like Naïve Bayes and Random forest algorithm are used for the prediction. These algorithms independently predict the results based on various parameters. This technique will also helps to upgrade the skills of students to meet the eligibility criteria.

Keywords: Random forest, Naïve Bayes Algorithm, Prediction, Machine learning

I. INTRODUCTION
The main success of an educational institution is measured using the percentage of placement. This system predicts about the probability of students to get placed. We have to find the best machine learning algorithm with maximum accuracy for our dataset. The placement cell can analyze their students graph based on the practice tests. This helps the students to realize which area they should work upon more. The process will be very effective for students as well as placement cell to get a good placement record.

II. PROBLEM STATEMENT
We cannot consider the placement of students just by their academic performances because some students may be good at aptitude, technical and communication skills. Due to their low academic score may tend to their drawback. For prediction process we need parameters like CGPA, logical and technical skills etc.

III. METHODOLOGY
WEKA is a data mining tool written in JAVA which is used to classify the accuracy of different algorithms using different datasets.

(1) Explorer: This interface has different panels like preprocess, associate, classify, cluster, select attribute and visualize.

(2) Experimenter: This interface provides facility for systematic comparison of different algorithms on basis of given datasets.

Fig 1 visualizes the interfaces of WEKA data mining tool.

IV. DATASETS
Dataset is a data collection or a single statistical data where every attribute of data represents a variable and each instance has its own description. For prediction process we used student data set and classification of algorithms in order to compare their accuracy using WEKA’s interfaces.
V. DATA MINING TECHNIQUES

Data mining techniques are used to predict the placement of students. In this paper we are using WEKA data mining tool for Predictions. Here classifications and accuracy are done by applying various algorithms.

A. NAÏVE BAYES CLASSIFIER

Naïve Bayes classifier assumes that the presence of a particular feature is unrelated to the presence of any other feature. The performance of Naïve Bayes is very effective on real data applications. It helps to build the fast machine learning models to make predictions quickly.

B. WORKING OF NAÏVE BAYES

[1] Click on Explorer tab
[2] Click on open file and select the dataset
[3] Click on classify tab
[4] Click on choose tab and select Naïve Bayes.

C. RANDOM FOREST

The random forest is a classification algorithm which is used for both classification as well as regression. It creates decision trees on data samples and then gets prediction on each of them and finally selects the best solution.

D. WORKING OF RANDOM FOREST ALGORITHM

[1] Click on Explorer tab
[2] Click on open file and select the dataset
[3] Click on classify tab
[4] Click on choose tab and select Random forest.

VI. RESULTS

![Confusion Matrix for Naïve Bayes](image)

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>Classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>148</td>
<td>0</td>
<td>a = Placed</td>
</tr>
<tr>
<td>19</td>
<td>48</td>
<td>b = Not Placed</td>
</tr>
</tbody>
</table>

![Confusion Matrix for Random Forest](image)

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>Classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>114</td>
<td>27</td>
<td>a = No</td>
</tr>
<tr>
<td>57</td>
<td>17</td>
<td>b = Yes</td>
</tr>
</tbody>
</table>

VII. CONCLUSION

Predicting student placement class manually by a placement officer is a difficult task. To resolve this, we can use data mining to predict the student placement. Our proposed system implements a student placement prediction system which predicts a particular student is placed or not with the help of algorithms Naïve Bayes and Random Forest. The result shows that Naïve Bayes has the highest accuracy for prediction. Using the results it is clear that the student dataset containing placement and academic details are essential source for predicting the future placement chances of the student.
VIII. REFERENCES

[1] D. Satish Kumar, Zailan Bin siri, D. S Rao, Anusha “Predicting student’s campus placement probability using binary logistic regression” 2019


