Campus Placements Prediction & Analysis using Machine Learning

ABSTRACT

Placement of students is one of the most important objective of an educational institution. Reputation and yearly admissions of an institution invariably depend on the placements it provides it students with. That is why all the institutions, arduously, strive to strengthen their placement department so as to improve their institution on a whole. Any assistance in this particular area will have a positive impact on an institution’s ability to place its students. This will always be helpful to both the students, as well as the institution. In this study, the objective is to analyse previous year's student's data and use it to predict the placement chance of the current students. This model is proposed with an algorithm to predict the same. Data pertaining to the study were collected form the same institution for which the placement prediction is done and also suitable data pre-processing methods were applied. This proposed model is also compared with other traditional classification algorithms such as Decision tree and Random forest with respect to accuracy, precision and recall. From the results obtained it is found that the proposed algorithm performs significantly better in comparison with the other algorithms mentioned.

**EXISTING SYSTEM**

Ajay Shiv Sharma, Swaraj Prince, Shubham Kapoor, Keshav Kumar used Logistic regression technique on their college placement dataset which got 83.33% of accuracy [2]. Jai Ruby, Dr. K. David used ID3, J48, REP Tree, NB Tree, MLP, Decision Table Classification techniques on the placement dataset collected from their college. The results had shown that ID3 predicted well among them with an accuracy of 82.1% [3]. Ankita A Nichat, Dr. Anjali B Raut used C4.5 classification technique on the placement dataset which was collected from their college which got 80% of accuracy [4].

NaiveBayes, OneR classification techniques on the data gathered from their high school. The results had shown that J48 and Simple Cart predicted well among them with an accuracy of 79.61% [5]. Ajay Kumar Pal and Saurabh Pal collected the data for the study and analysis of the student’s educational performance basically for training and placement. The authors used different classification algorithm and used WEKA data mining tool [6]. They concluded that naive Bayes classification model is the better algorithm based on the placement data with found accuracy of 86.15% and overall time taken to build the model is at 0 sec. As compared with others Naïve Bayes classifier had lowest average error i.e. 0.28.

Ravi Tiwari and Awadhesh Kumar Sharma built the prediction model to improve the placement of the students [7]. They used WEKA as the data mining tool to build the model using random tree algorithm. They also used ID3, Bayes Net, RBF network, J48, algorithms on the student data set. They resolved that the RT (Random Tree) algorithm is more accurate with 73% for the classification/prediction of the model. The accuracy using ID3 and J48 is 71%. Bayes Net is 70%

**Disadvantages**

* The system is not implemented Attribute selection which is not relevant to each other.
* The system is not implemented Cleaning missing values.

Proposed System

• In Placement Prediction system predicts the probability of a undergrad students getting placed in a company by applying classification algorithms such as Decision tree and Random forest. The main objective of this model is to predict whether the student he/she gets placed or not in campus recruitment. For this the data consider is the academic history of student like overall percentage, backlogs, credits. The algorithms are applied on the previous years data of the students.

**Advantages**

* Some of the attributes in the initial dataset that was not pertinent (relevant) to the experiment goal were ignored. The attributes name, roll no, credits, backlogs, whether placed or not, b.tech % ,gender are not used.
* The random forest algorithm can also be thought of as an ensemble method in machine learning. The input to a random forest algorithm is a dataset consisting of records, with attributes. Random subsets of the input are created.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).