

# STUDENT'S PLACEMENT PREDICTION SYSTEM

\* Khamkar Pratiksha, \*\*Lagad Rutuja, \*\*\*Shinde Priyanka, \*\*\*\*Londhe Shubhangi.

Project Guide: Prof.Bhosle Swati.S.

Department of Computer Engineering,

HSBPVT'S Group of Institution's, College of Engineering, kashti, A.nagar, Maharashtra, India.

## Abstract

*Placement of students is one in every of the vital activities in academic establishments. Admission and name of establishments primarily depends on placements. Hence all institutions strive to strengthen placement department. The main objective of this paper is to analyze previous year's student's historical data and predict placement possibilities of current students and aids to increase the placement percentage of the institutions. We are not going to consider the placement of students not only by their academic performances but also aptitude, technical and communication skills, and segregating total student placement data into students placed in different streams to identify in which stream placements are more and use that data to predict the next year admission trends. Here we use different machine learning classification algorithms, namely KNearest Neighbors [KNN] algorithm, Random Forest. These algorithms independently predict the results and we then compare the efficiency of the algorithms, which is based on the dataset. This model helps the position cell at intervals a corporation to spot the potential students and concentrate to and improve their technical and social skills.*

## Keywords:

Student Placement Prediction, KNN Algorithm, Random forest.

---

## I. INTRODUCTION:

We aim to develop a placement predictor as a part of making a placement management system at college level which predicts the probability of students getting placed. It will also help the teachers as well as placement cell in an institution to provide proper care towards the improvement of students in the duration of course. We are using machine learning for the placement prediction. The existing placement prediction model considers only academic performances of the students so that the prediction of the student getting placed or not can be done. 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 score in their academic that may tend to be their drawback. For predicting the placement of a Student needs parameters like cgpa, logical and technical skills Academic performances may be important but the model is design to predict the placements based on the parameters of the students and segregating the total students placed data according to their streams to find out that in which stream placements are more. Using this model we can predict the next year admission trends.

## II. LITERATURE REVIEW:

Wilton W.T. FOK, Y.S. He, H.H Au Yeung and K.Y. Law conducted a study to predict suitable course for the students, based on their behavior using Neural Network Technique. Tensor Flow engine includes number of intermediate node and number of deep learning layers are adjusted and compared [1].

Dr. A. Padmapriya\* (November 2012) Data mining algorithm Decision Tree Induction is best when compared to Naive Bayesian Classifier according to Classification Accuracy misclassification Rate, Speed & Size on students Personal Data, Pre college data & Under Graduation Data to predict Higher Education Admissibility [2]

H. Sabnani, M. More, P. Kudale, S. Janrao, (2018) proposed Prediction of Student Enrolment Using Data Mining Techniques. They have used the Apriori technique to analyze the behavior of students who are seeking admission to a particular college. They have also used the Naïve Bayes algorithm which will help students to choose the course and help them in the admission procedure. In their project, they were conducting a test for students who were seeking admissions and then based on their performance, they were suggesting students a course branch using Naïve Bayes Algorithm [3]

Animesh G. al., 2016 have proposed a system to predict the probability of a student being placed using k nearest neighbour classifier. Euclidean Distance is used as a distance measure. The dataset involves academic history of a student such as 10th percentage, 12th percentage, aggregate of engineering scores and number of backlogs. Further post processing is done using scores of Technical skills, Communication skills, Analytical Skills and as well as Teamwork scores are used [4]

Karan P. and Parteek B., 2015 have presented a paper on comparative study of different data mining techniques such as J48 decision tree algorithm, Naive Bayes for predicting a student's placement in different types of companies. The attributes of the dataset include cumulative grade point, different subject scores and the type of company. Missing data values were replaced with mean value for numeric data and mode value for nominal data [5].

Patel T., et al "Data Mining Techniques for Campus Placement Prediction in Higher Education." Indian J.Sci.Res. 14 (2) 2017. In this paper, the creator had driven evaluations on the utilization of information digging frameworks for grounds position supposition and use of WEKA programming for plan and execution. Different cutoff points which could be considered for figuring understudy execution are the scholastic show, social limits, specific limits, capable plan and tries. Different pressing assessments like clear k-mean, Farthest-first convergence, segregated assembling, moderate grouping were used for model turn of events. It was seen that the time is taken for building clear k-mean, Farthest-first intersection point and bound gathering was just 0.02sec in regards to various evened out squeezing (0.09 sec) and thickness based collecting (0.08 sec)[6].

### III. EXISTING SYSTEM:

The existing placement prediction model considers only academic performances of the students so that the prediction of the student getting placed or not can be done.

#### DRAWBACK:

- In existing model only academic performances of student's are considered and no other details are taken into consideration like internship, aptitude, skills, etc. .

### IV. PROPOSED SYSTEM:

In proposed system, we are not considering 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 score in their academic that may tend to be their drawback. For predicting the placement of a Student needs parameters like cgpa, logical and technical skills Academic performances may be important but the model is design to predict the placements based on the parameters of the students and segregating the total students placed data according to their streams to find out that in which stream placements are more. Using this model we can predict the next year admission trends.

#### ADVANTAGES:

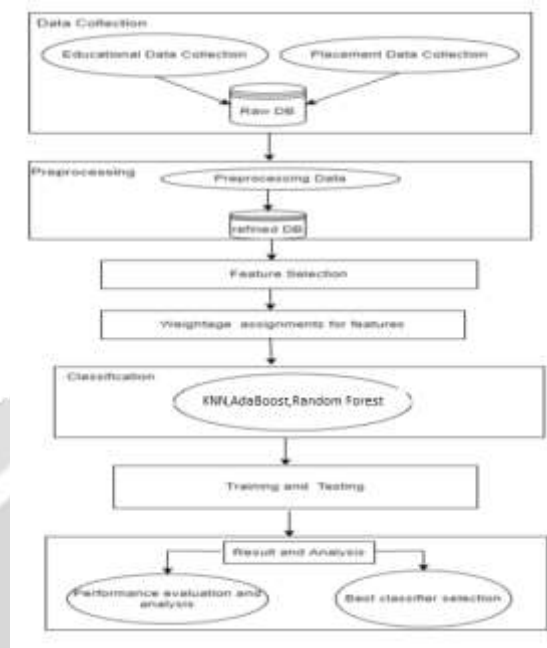
- Proposed model can predict the placement more accurately because of we are using student's technical skills, communication skills, and also the internships done by the students in this predictive model.
- We can use this model to predict the next year admission trends.

### V. PROPOSED METHODOLOGY:

#### Algorithms:

- **KNN Algorithm:**

KNN stands for K-nearest neighbor, it's one of the supervised learning algorithm mostly used for classification of data on the basis how its neighbor are classified. KNN stores all available cases and classifies new cases based on a similarity measure. K in KNN is a parameter that refers to the number of the nearest neighbors to include in the majority voting process.



**Fig1: Block Diagram of Methodology**

- **How does KNN work?**

The K-NN working can be explained on the basis of the below algorithm:

- Step-1: Select the number K of the neighbors
- Step-2: Calculate the Euclidean distance of K number of neighbors
- Step-3: Take the K nearest neighbors as per the calculated Euclidean distance.
- Step-4: Among these k neighbors, count the number of the data points in each category.
- Step-5: Assign the new data points to that category for which the number of the neighbor is maximum.
- Step-6: Our model is ready.

- **Random Forest Algorithm**

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

**How does Random Forest algorithm work?**

Random Forest works in two-phase first is to create the random forest by combining N decision tree, and second is to make predictions for each tree created in the first phase.

The Working process can be explained in the below steps and diagram:

- Step-1: Select random K data points from the training set.
- Step-2: Build the decision trees associated with the selected data points (Subsets).

Step-3: Choose the number N for decision trees that you want to build.

Step-4: Repeat Step 1 & 2.

Step-5: For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

## VI. RESULTS:

### 1. Random Forest :

The final accuracy obtained for the test data set using the Random Forest model was **87.46%**.

### 2. KNN:

To reach the maximum accuracy with this model the k value is 6. The model is doing a great job with a high accuracy and only not predicting very few "Placed" student. We could achieve higher precision of 100% with k > 14 but we could be suffering from overfitting our data. Thus the final accuracy obtained is **~85%**

## VII. FUTURE ENHANCEMENT

The future enhancements of the project is to focus on to add some more parameters to predict more efficient placement status. We can also enhance the project by predicting some solutions or suggestions for the output generated by system.

## VIII. CONCLUSION:

Student Placement Predictor is a system which predicts student placement status using machine learning and in which stream placements are more to predict the next year admission trends. We can use this predictions for counseling the students and their parents, which stream has a higher scope in future.

This paper also proposed a methodology to predict the placements of candidates. KNN, Random Forest models were developed as a part of the solution. The Random Forest was found to have an accuracy of 88.41% when compared to the KNN. Thus it can be concluded that the Random Forest model is the best fit for the current scenario.

## IX. REFERENCES:

1. "Prediction Model for Students Future Development by Deep Learning and TensorFlow Artificial Intelligence Engine" 2018 4th IEEE International Conference on Information Management
2. Dr. A. Padmapriya\* (November 2012) Prediction Of Higher Education Admissibility Using Classification Algorithms International Journal of Advanced Research in Computer Science and Software Engineering ISSN: 2277 128X
3. H. Sabnani, M. More, P. Kudale, S. Janrao, "Prediction of Student Enrolment Using Data Mining Techniques", International Research Journal of Engineering and Technology (IRJET), 5(4), 1830-1833, 2018.
4. Animesh, G., M. Vignesh., Bysani, P., Naini, D., 2015. A Placement Prediction System Using K-Nearest Neighbors Classifier. 2nd International Conference on Cognitive Computing and Information Processing (CCIP).
5. Karan. P., Prateek, B., 2015. "Application of Data mining in predicting student placement", International conference on Green computing and Internet of things (ICGCIoT).
6. Patel, T., Tamrakar, A. "Data Mining Techniques for Campus Placement Prediction in Higher Education." Indian J.Sci.Res. 14 (2) 2017: 467-471.

## X. APPENDIX VISUALIZATION:



Fig1: Home Page

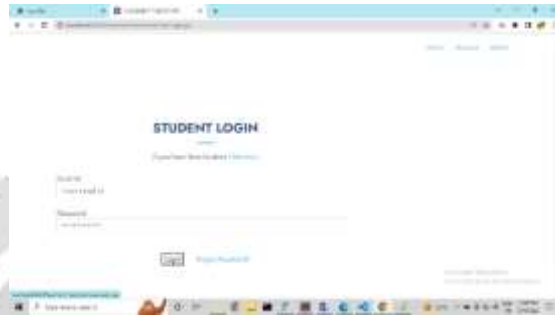


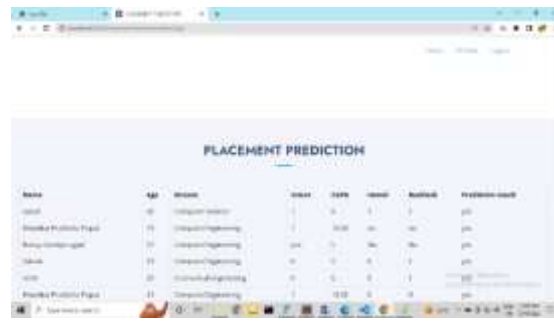
Fig2: Student Login



Fig3: Admin Login



Fig 4:Test Data



Name	Age	Marks	Placement Rate	Placement Rate	Placement Rate	Placement Rate	Placement Rate
John	20	85	90	95	98	99	100
David	21	78	82	88	92	95	98
Emily	22	88	92	96	99	100	100
Michael	23	75	80	85	90	95	98
Anna	24	82	88	92	96	99	100
James	25	70	75	80	85	90	95
Sarah	26	80	85	90	95	98	100

Fig5: Placement Prediction

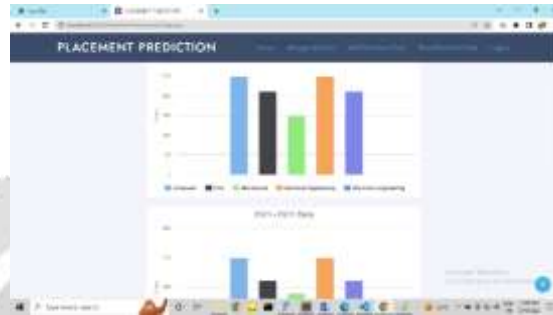


Fig:Placement Graph

