iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.59601

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

A Placement Prediction System for Polytechnic Students

Parvathy G S[1], Smitha R[2]
[1]*Dept. of Computer Engineering, NSS Polytechnic College, Pandalam, Kerala, India*
[2]*Dept. of Electronics and Communication Engineering, NSS Polytechnic College, Pandalam, Kerala, India*

Abstract: ***The placement prediction system makes predictions about a candidate's likelihood of being hired based on several factors, including talents, backlogs, CGPA, and more. Here, previous placement information is analyzed to identify success parameters and create a machine-learning model that forecasts placement results in the future. It is designed to encourage students to improve their academic performance, enhance their skill set, and develop additional soft skills to increase their likelihood of securing successful job placements. This system guarantees that educational establishments are adaptable to the changing demands of the job market and can proficiently equip learners for prosperous career prospects. This study aims to predict the placement outcomes of polytechnic students in Kerala using three different classification algorithms: K-nearest neighbors (KNN), Gaussian Naïve Bayes Classifier, and Support Vector Classifier (SVC). By employing these algorithms, the most effective method for predicting whether a student will be placed or not after completing their polytechnic education is identified. Student's CGPA and scores obtained from skill tests have been utilized for placement prediction.***

Keywords: Machine learning (ML), K-Nearest Neighbors (KNN), Random Forest Classifier, Support Vector Classifier (SVC), Gaussian Naïve Bayes Classifier, Deep Neural Network, Iterative Dichotomiser 3 (ID3)algorithm

I. INTRODUCTION

The ultimate aim for students is to secure their ideal job. However, due to ignorance about the precise placement criteria, achieving this goal can be difficult. Placement prediction systems can be used to forecast a student's chances of securing a job or employment. The system enables students to make well-informed decisions about their educational and professional tracks by offering insights into the aspects that influence placement outcomes, including academic achievements, skill competency, and the development of soft skills. It analyses a variety of aspects influencing the effectiveness of job placement, including CGPA, skills, and other criteria, by utilizing statistical analysis, machine learning algorithms, and predictive modeling approaches [1][2]. Employers, educational institutions, and students can all benefit from using Placement Prediction Systems, which are computer solutions that help forecast and optimize job placement outcomes.

Three machine learning techniques are used by the system: Gaussian NB, Support Vector Classifier (SVC), and K-Nearest Neighbours (KNN). The placement prediction method makes use of information gathered from questionnaires given to polytechnic students across different departments. These questionnaires are designed to gather comprehensive information about the student's CGPA, skills, gender, branch, etc. A skill test (English proficiency, technical skill, mathematical skill) has been conducted on students and the score is evaluated. Heat map, the correlation with the target variables has also been presented in the system for better understanding.

II. REVIEW ON EXISTING APPROACHES

The decision tree and random forest algorithms were applied in [3]to categorize a dataset of students who were placed on campus and those who were not. Out of these two, the random forest prediction model achieves 86% accuracy. As a result of the random forest's ensemble nature, it can synthesize weak predictions into an overall forecast that is more reliable and accurate. A student placement possibility prediction utilizing the Naive Bayes algorithm has been proposed in [4]. Naive Bayes algorithms are well-suited for classification jobs with many features since they rely on feature independence and follow Bayesian principles. This system analyzes three different Bayes classifiers: the Multinomial Naïve Bayes, Bernoulli Naïve Bayes, and Bernoulli Naïve Bayes algorithms.

©IJRASET: All Rights are Reserved

The results reveal that the Bernoulli Naïve Bayes algorithm has the highest accuracy, achieving an astounding 75%. KNN and Euclidian metrics have been utilized in [5] for student placement prediction. A logistic regression model was used to develop a placement predictor system (PPS) in [6]. An open-source GNU Octave programming tool was used to produce the results. The model utilizes student scores in matriculation, senior secondary, and subjects in various semesters of technical education, and demographics for prediction tasks. With an accuracy percentage of 83.33%, the Placement Prediction System (PPS) testing has produced favorable results. Deep Neural Network (DNN) has been used for students' placement prediction tasks in [7]. The depth of DNNs enables them to recognize complex patterns in the input and learn significant attributes of the data. This prediction model gives 97% accuracy for placement prediction. ID3 algorithm has been used for placement prediction in [8]. The ID3 (Iterative Dichotomiser 3) classification technique, based on the decision tree algorithm, has been employed in the context of classification tasks.

III. PROPOSED SYSTEM

A placement predictor system for predicting the placement result possibility of polytechnic students in Kerala has been proposed here. To determine whether a student will be placed or not, the current placement prediction algorithm simply takes into account their academic performance. But here, a skill test has been conducted on students and the obtained score along with other variables has been utilized. Fig. 1 shows the block diagram of the placement prediction system. The system includes

- Dataset preparation
- Feature Selection
- Model Selectio
- Model training and valuation

A. Dataset Preparation & Feature Selection

The first step towards developing a placement prediction system involves gathering necessary information from students. Students at Polytechnic College are first asked to fill out Google forms with appropriate information sources, such as their academic records, skill tests, demographic data, and previous placement data.

After that, the data collected is cleaned and given a preprocessing step to address any missing values, unusual values, or inconsi stencies. It is necessary to choose the most relevant characteristics that are indicative of placement results. One may use feature selection methods like dimensionality reduction, correlation analysis, or feature importance ranking.

B. Model Selection

For the placement prediction task, suitable machine learning algorithms or predictive models are selected. Decision trees, logistic regression, support vector machines, neural networks, and ensemble techniques like random forests or gradient boosting are examples of common models. In this system, KNN, SVC, and Gaussian NB algorithms are used for classification tasks.

1) *KNN Classifier:* The KNN classifier works by keeping the whole training dataset in memory and predicting new, unlabelled instances by comparing them to previously classified examples[9]. For classification, KNN uses Euclidean distance, to calculate the distance between each new data point and every data point in the training dataset. It then chooses the closest neighbors by figuring out how far away the new instance is from K's nearest neighbors. A hyperparameter called K decides the number of neighbors considered during classification.

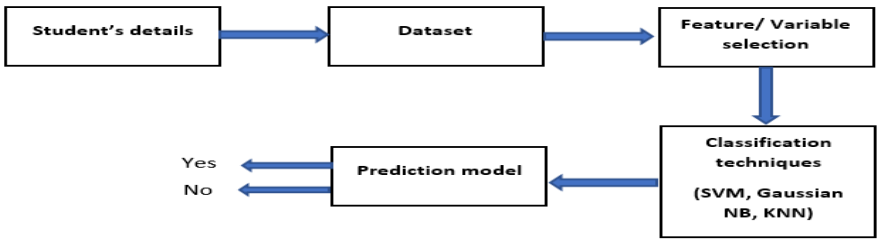


Fig.1: Placement prediction system block diagram

2) *SVC Classifier:* SVC is a powerful technique that works well for binary classification applications [10]. It can deal with data that is *both* linearly and non-linearly separable. The goal of this approach is to identify the optimal hyperplane for classifying a feature space. Hyperplanes are selected to have the largest potential margin—that is, the distance between them and the closest data points. After being trained, the SVC model can be utilized to anticipate the location of previously undiscovered instances by determining which side of the hyperplane they reside on. One class is assigned to instances on one side of the hyperplane, and a different class is assigned to those on the opposite side.
3) *Gaussian NB:* A variation of the Naive Bayes technique which is capable of data classification with continuous or real-valued information is the Gaussian Naive Bayes (GNB) classifier [11]. According to GNB, each category's features exhibit a Gaussian distribution. For training the classifier the mean as well as the variance of every feature are computed for every class in the training dataset. Using the Gaussian PDF, GNB predicts the probability that a given subject falls into each of the categories.

C. Model Training and Valuation

Machine learning models have been evaluated for performance and their ability to generalize to new, unknown data using the cross-validation technique. Cross-validation is used to train a machine-learning model on pre-processed data. Models modify their parameters throughout training to lower prediction errors. Following the models' training and fine-tuning, their performance is evaluated using metrics derived from cross-validation.

IV. EXPERIMENTAL RESULTS

A placement prediction system has been developed using student data gathered from NSS Polytechnic College students across multiple branches. Cumulative Grade points, Skill test scores, 10th standard board, demographic information (sex, branch, and age), etc. have been collected from students. On training and testing the three classification models (SVC, KNN, Gaussian NB), upon cross-validation, the Gaussian NB classifier gives the best validation score. The average of the evaluation scores acquired over all iterations is usually the cross-validation score. A model's ability to generalize is indicated by a higher cross-validation score. The validation scores of the three models are shown in Fig 2. The accuracy of placement prediction for the Gaussian model is 95.45%. Only 1 wrong prediction has been made out of 22 cases. Training and testing accuracy of the Gaussian NB prediction model has been shown in Fig.3. A heat map is a graphical representation used in machine learning that uses colours to show how characteristics or variables in a dataset relate to one another. Heat maps are frequently used to find clusters, relationships, or patterns in data. Every cell in the heat map is given a colour according to the correlation coefficient it indicates. The correlation between the variables represented by the heat map is depicted in Fig.4.It is found that CGPA and skill test score are more correlated with the target variable 'placement status'.10th Board does not have any association with placement status.

```
_____________________________________________________
cross validation model : KNeighborsClassifier
Mean score :  88.9090909090909
_____________________________________________________
cross validation model : SVC
Mean score :  91.54545454545456
_____________________________________________________
cross validation model : GaussianNB
Mean score :  92.54545454545455
```

Fig.2 : Validation scores of three models

```
model trained with GaussianNB()
Model accuracy on train is::  0.9285714285714286
Model accuracy on test is::  0.9545454545454546
Confusion matrix train is::  [[38  2]
 [ 4 40]]
Confusion matrix test is::  [[12  1]
 [ 0  9]]
Wrong Predictions made: 1 / 22
```

Fig.3 Training and testing accuracy of Gaussian NB model

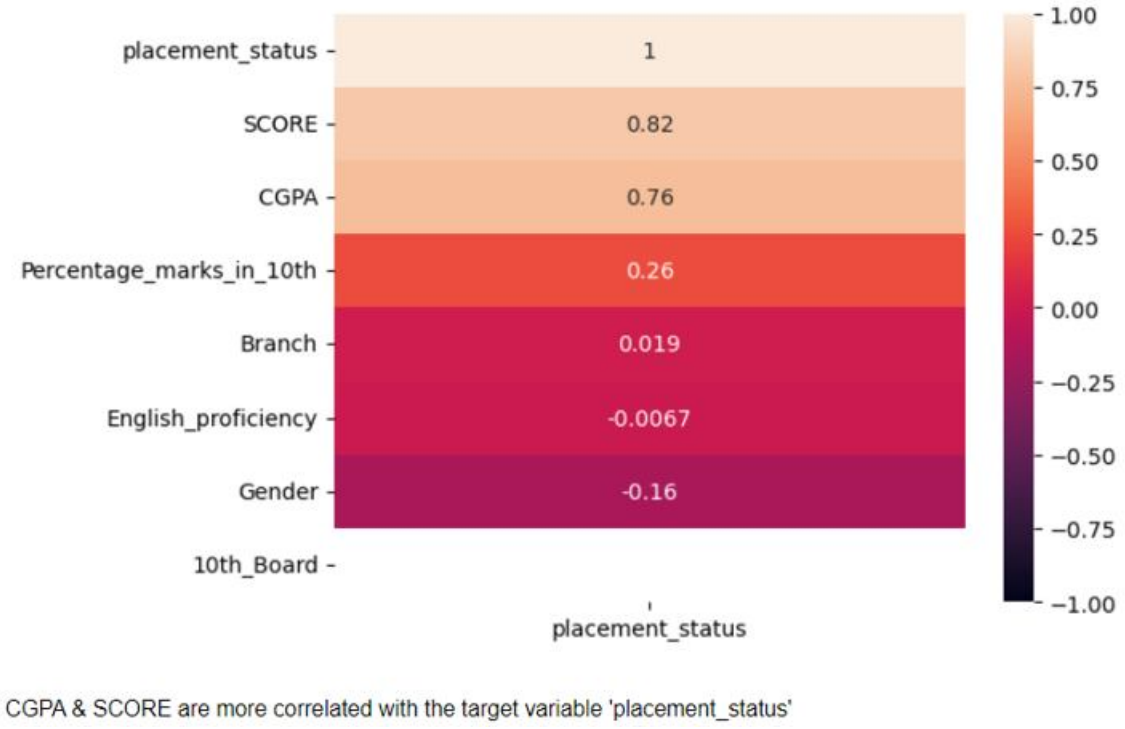


Fig.4: Heat Map

V. CONCLUSION

Student placement is one of the most significant responsibilities of educational institutions. Finding the students who require additional assistance and implementing the right strategies to improve their skills are crucial steps toward getting a successful placement. A placement prediction system based on machine learning models has been introduced in this paper. Extensive research on three placement prediction models has yielded insightful findings. For colleges and universities, this predictive method is a useful tool for efficient planning and strategy. Three machine learning models named Gaussian NB, KNN, and SVC classifiers have been cross-validated and it is found that the best prediction model is the Gaussian NB model. The Gaussian model performs placement prediction with a classification accuracy of 95.45% accuracy. It is now clear that the student dataset, which includes academic and placement information, has enormous potential for use as a source for precisely projecting future placement prospects.

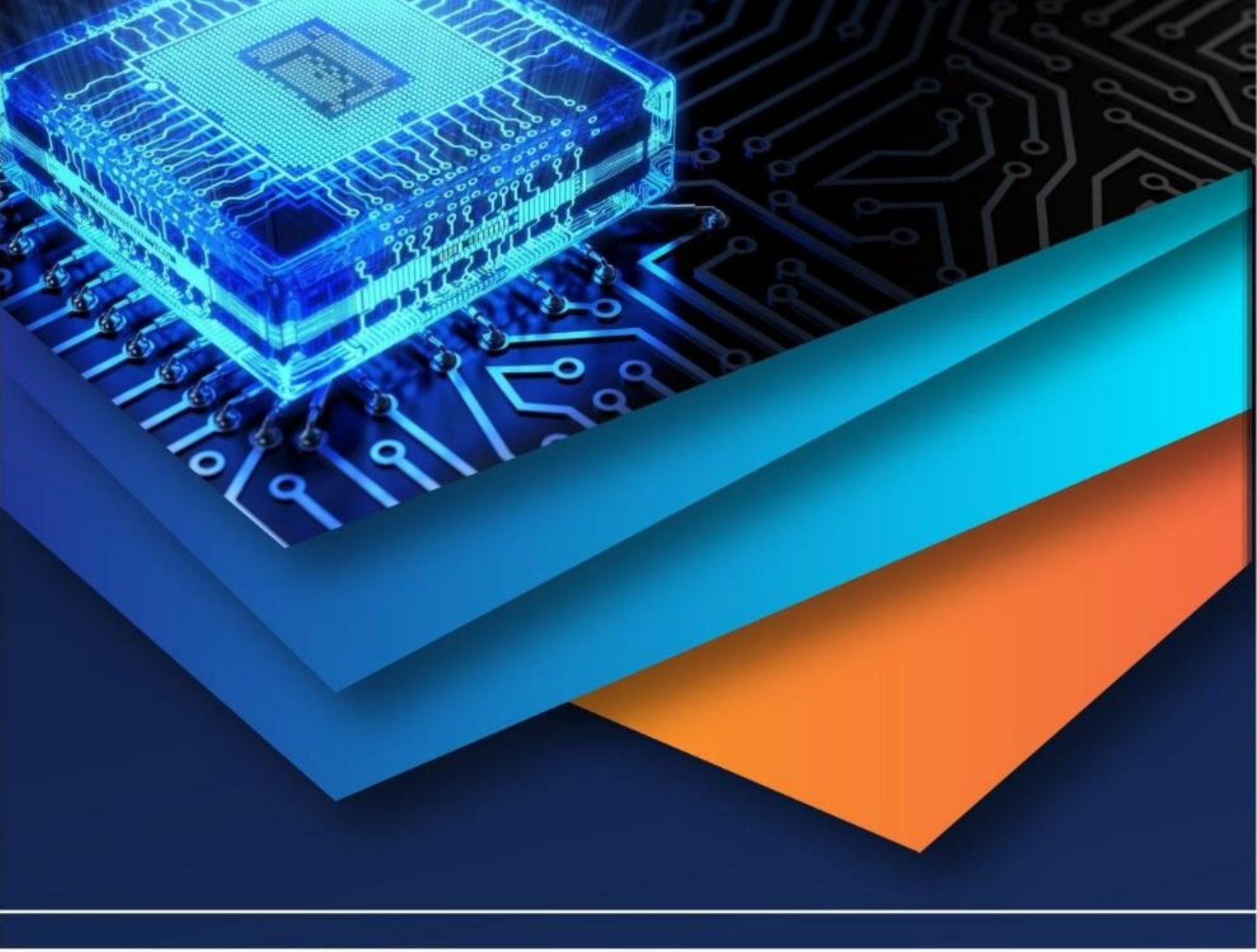
A. Conflict of Interest

The authors have no conflicts of interest to declare.

REFERENCES

[1] Joy, L. C., & Raj, A. (2019, March). A review on student placement chance prediction. In 2019 5th International conference on advanced computing & communication systems (ICACCS) (pp. 542-545). IEEE.
[2] Thangavel, S. K., Bkaratki, P. D., & Sankar, A. (2017, January). Student placement analyzer: A recommendation system using machine learning. In 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 1-5). IEEE.
[3] Manvitha, Pothuganti, and Neelam Swaroopa. "Campus placement prediction using supervised machine learning techniques." Int J Appl Eng Res 14.9 (2019)
[4] Brijmohan Lal Sahu , Dr. Anil Tiwari, “Student placement possibility prediction using Naive Bayes algorithm” International Journal of Advance Research, Ideas and Innovations in Technology, ISSN: 2454-132X Impact factor: 6.078 (Volume 6, Issue 3)
[5] 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)
[6] Sharma, Ajay & Prince, Swaraj & Kapoor, Shubham & Kumar, Keshav. (2015). PPS - Placement prediction system using logistic regression. 337-341. 10.1109/MITE.2014.7020299.

[7] Udawat, B., Kale, A., Sinha, D., Sharma, H., Krishnan, D. (2022). Campus Placement Prediction System Using Deep Neural Networks. In: Khanna, A., Gupta, D., Bhattacharyya, S., Hassanien, A.E., Anand, S., Jaiswal, A. (eds) International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing, vol 1394. Springer, Singapore. https://doi.org/10.1007/978-981-16-3071-2_3
[8] Bhatt, H., Mehta, S., & D'Mello, L. (2015). Use of ID 3 Decision Tree Algorithm for Placement Prediction.
[9] https://www.geeksforgeeks.org/k-nearest-neighbours/
[10] https://www.analyticsvidhya.com/blog/2021/01/a-quick-introduction-to-k-nearest-neighbor-knn-classification-using-python/
[11] https://www.geeksforgeeks.org/gaussian-naive-bayes-using-sklearn/
[12] https://scikit-learn.org/stable/modules/svm.html
[13] https://www.tutorialspoint.com/scikit_learn/scikit_learn_support_vector_machines.htm
[14] https://www.tutorialspoint.com/scikit_learn/scikit_learn_gaussian_naive_bayes.htm
[15] https://dataaspirant.com/gaussian-naive-bayes-classifier-implementation-python/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)