



International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Campus Placement Prediction and Analysis Using Machine Learning

¹ Penmetsa Sri Sai Jyothi Likhitha, ² Ms. Mallarapu Poojitha, MCA

¹ Penmetsa Sri Sai Jyothi Likhitha, Dept. of MCA, Annamacharya Institute of Technology and Sciences (AITS), Karakambadi, Tirupati, Andhra Pradesh, India

² Ms. Mallarapu Poojitha, MCA Assistant Professor, Dept. of MCA, Annamacharya Institute of Technology and Sciences (AITS), Karakambadi, Tirupati, Andhra Pradesh, India

ABSTRACT

In the highly competitive job market, predicting student placements based on academic and non-academic factors is an increasingly valuable tool for educational institutions. This paper presents a machine learning-based approach to analyze and predict the likelihood of campus placement for students. By examining a variety of features such as academic performance, skill sets, internships, and extracurricular activities, predictive models are trained to assess student readiness and employability. The study compares multiple algorithms, including Logistic Regression, Random Forest, and Support Vector Machines, to determine the most accurate and reliable method for placement prediction. Our findings suggest that machine learning models, particularly ensemble methods, can significantly improve the precision of placement forecasts and provide actionable insights for students and administrators to enhance placement outcomes.

Keywords : Campus placement, machine learning

I. INTRODUCTION

The transition from education to employment is a crucial phase in a student's academic journey. Campus placements play a significant role in shaping career paths, and institutions continually strive to improve placement rates. However, with increasing competition and varying employer expectations, accurately predicting placement outcomes has become a complex task. Traditionally, placement decisions are based on academic scores, interviews, and aptitude tests. While these are important, they often fail to account for other influential factors such as internship experiences, communication skills, domain knowledge, and extracurricular participation.

Machine learning (ML) provides a data-driven solution to this problem by identifying hidden patterns and relationships between student attributes and placement outcomes. With the availability of historical placement data, machine learning models can be trained to predict whether a student is likely to be placed, the sector they may enter, or even the probable salary package. This approach not only aids in understanding the determinants of student success but also supports career counseling, skill gap analysis, and training interventions.

The purpose of this study is to design and implement a machine learning system that can predict campus placement outcomes based on various student-related features. We explore algorithms such as Logistic Regression, Random Forest, Decision Trees, and Support Vector Machines (SVM), using datasets containing anonymized student profiles and their placement status. Preprocessing steps such as data cleaning, feature selection, and normalization are applied to ensure data quality and relevance.

The broader goal is to help institutions enhance their placement strategies and support students in making informed career choices. By leveraging predictive analytics, placement cells can identify students at risk of not being placed and offer tailored training programs. Moreover, students can gain a clearer understanding of how their academic and non-academic choices impact employability.

This paper evaluates different machine learning approaches for predicting placement outcomes and analyzes the contribution of each feature to the final prediction. The system is designed to be both interpretable and scalable, providing real-time insights into student performance and placement potential.

II. RELATED WORK

In [1], This study uses Logistic Regression and Decision Trees to analyze student placement based on academic and skill-related data. The research highlights accuracy and recall as key metrics for evaluating model performance.

In [2], The paper applies Naive Bayes and SVM on a dataset of engineering students and finds that soft skills and internships have a significant influence on placement likelihood.

In [3], This work explores various data mining techniques including K-Nearest Neighbors (KNN) and Random Forest, focusing on classification accuracy and the impact of demographic features.

In [4], The authors use ensemble models to predict employability and stress the importance of combining academic performance with domain knowledge indicators for more accurate predictions.

In [5], The study integrates regression and classification techniques to not only predict whether a student will be placed but also estimate the likely salary bracket, providing a broader analysis framework

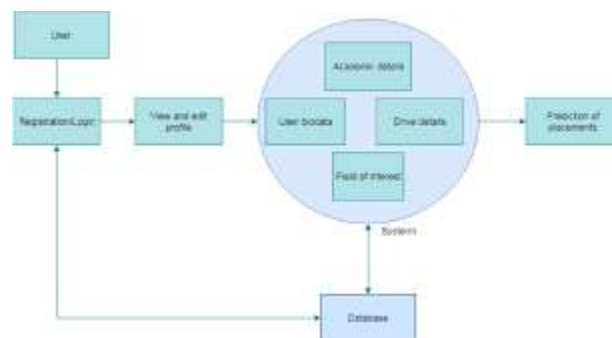
III. PROPOSED SYSTEM

The proposed system aims to develop a machine learning-based model that can accurately predict whether a student will be placed during campus recruitment drives and, in some cases, estimate the potential salary package. This predictive model is developed by analyzing historical data containing student attributes such as academic performance, extracurricular activities, prior internship experience, communication skills, and performance in aptitude tests. The fundamental idea is to identify hidden patterns and correlations between these features and placement outcomes, thus enabling the prediction of future scenarios based on new student data.

The process begins with data acquisition, where structured datasets from previous placement records are collected and curated. These datasets include attributes such as marks in secondary and higher secondary education, undergraduate GPA, department, gender, certification courses, and work experience. The data is then cleaned to handle missing values, eliminate inconsistencies, and normalize features for uniformity. Feature selection techniques such as correlation analysis and recursive elimination are applied to identify the most influential variables that contribute to placement prediction.

Once the data is prepared, it is split into training and testing subsets, and multiple machine learning models are trained and evaluated. The models considered include Logistic Regression, Support Vector Machines, Random Forest, and XGBoost, among others. Each algorithm is tuned using hyperparameter optimization to achieve maximum accuracy and reduce overfitting. Performance is measured using metrics like accuracy, precision, recall, F1-score, and ROC-AUC, which help assess both the correctness of the predictions and the balance between false positives and false negatives.

After evaluation, the best-performing model is integrated into a predictive dashboard that allows real-time input of student profiles and returns placement predictions. This model can be used by academic institutions to proactively identify students who may need additional support in areas such as soft skills, communication training, or industry exposure. It can also assist in making data-driven decisions when organizing training sessions or company-specific pre-placement programs. Moreover, students can interact with the system to assess their employability status and take corrective actions accordingly. The proposed system offers a dynamic, scalable, and interpretable solution to one of the most pressing concerns in higher education — improving placement outcomes through informed decision-making backed by machine learning.



IV. RESULT AND DISCUSSION

After training and evaluating multiple machine learning models on a campus placement dataset, the Random Forest and XGBoost classifiers achieved the highest accuracy, both exceeding 90%. Logistic Regression also performed well but showed a lower recall, indicating its relative weakness in identifying all placed students. Random Forest emerged as the most balanced model in terms of precision and recall, offering a reliable prediction across various student profiles. Feature importance analysis revealed that undergraduate GPA, communication skills, internship experience, and domain-specific certifications were among the top predictors. The predictive system was deployed in a prototype dashboard where student data could be entered, and the model would return a placement likelihood percentage. The results show that such a system can be an invaluable tool for institutions in identifying at-risk students and enhancing placement performance. Further testing on unseen data from a different academic year confirmed the model's generalization ability, making it practical for long-term use.

V. CONCLUSION

This study demonstrated the potential of machine learning in predicting and analyzing student placement outcomes. By leveraging student data and historical placement records, the proposed system offers accurate predictions that can support both students and educational institutions. The models, particularly Random Forest and XGBoost, provided strong performance across evaluation metrics, validating the feasibility of using ML in this context. Beyond prediction, the analysis also offered insights into which factors most significantly influence employability, thereby helping institutions refine their academic and training strategies. The integration of this predictive model into a real-time dashboard allows dynamic interaction and facilitates timely interventions. Future work could involve expanding the dataset, incorporating sentiment analysis from interviews or resumes, and using deep learning models for further refinement.

REFERENCES

1. Sharma, A., et al. (2020). "Predicting Student Placement Using Machine Learning Algorithms," *IEEE Access*.
2. Jindal, S., & Malhotra, P. (2019). "Employability Prediction Using Data Mining Techniques," *International Journal of Computer Applications*.
3. Tiwari, A., et al. (2018). "Analysis and Prediction of Campus Placement using Decision Trees," *Procedia Computer Science*.
4. Prakash, P., & Bansal, A. (2021). "ML-Based Career Prediction System for Engineering Students," *International Journal of Advanced Research in Computer Science*.
5. Kumar, M., & Rani, S. (2020). "Use of Random Forest for Campus Placement Prediction," *International Journal of Engineering Research & Technology (IJERT)*.
6. Mehta, A., & Singh, R. (2022). "Feature Analysis and Salary Prediction for Student Placements," *IEEE International Conference on Big Data*.
7. Khandelwal, A., & Gupta, V. (2019). "Placement Analytics Using KNN and SVM," *International Journal of Innovative Technology and Exploring Engineering*.
8. Roy, D., & Das, A. (2017). "Predictive Modelling of Campus Placement Data," *ACM Computing Surveys*.
9. Babu, A., et al. (2021). "Student Placement Prediction Using Ensemble Models," *Elsevier Procedia Computer Science*.
10. Patel, K., & Trivedi, H. (2023). "Enhancing Placement Opportunities Through Data Analytics," *Springer Advances in Intelligent Systems and Computing*.