International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 11 | November - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

C2C - Predictive Analysis of Student Campus Placement

Prof. A.M. Chadchankar[1], Kajol Pawar[2], Vilas Dasalkar[3], Nihar Jachak[4], Prasad Kurale[5]

[1]Assistant Professor, Department of Computer Engineering, NBNSTIC, Pune, Maharashtra, India
[2345]Student, Department of Computer Engineering, NBNSTIC, Pune, Maharashtra, India

---

Abstract - ***Student placement holds undeniable significance in academic institutions, influencing admissions and overall university prestige. Universities strive to enhance their placement services, and this study focuses on analyzing recent student data to predict placement potential. The advanced predictive algorithm aims to optimize university effectiveness in student placement, benefiting both students and the institution. The methodology relies on a meticulously preprocessed dataset from the same institution, showing a significant increase in accuracy compared to established classification algorithms. The campus placement prediction system employs educational data mining and parameters from college records and tests, utilizing logistic regression for accurate predictions. The model enhances academic performance and soft skills, outperforming traditional algorithms and providing valuable insights through predictive analytics. This study offers a comprehensive perspective on predictive analytics in campus placement prediction.***

Key Words : ***Data Mining, Placement, Logistic Regression, KNN, Supervised Machine Learning, Accuracy, Prediction, Classification, Decision Tree, Random Forest, Data Analysis, Predictive Analytics, SVM, XGBoost.***

1.INTRODUCTION

Highlighting the historical importance of campus placements in the educational landscape, these events mark a crucial transition from academia to the workforce. In the face of a competitive job market, the need for more effective placement strategies becomes imperative. Specific challenges arise, such as aligning student skill sets with job requirements and accurately predicting successful placements.

Addressing these issues through predictive analysis holds significant potential benefits. Research objectives include the development of a predictive model to streamline the campus placement process, ultimately aiming to enhance efficiency and success rates. The paper's structure encompasses sections on data collection, methodology, results, discussion, and conclusion.

In the context of the expanding educational landscape, the pivotal role of student placement in shaping a university's reputation cannot be overstated. The research employs sophisticated algorithms such as RandomForestAL and SVMAL. to predict and analyze placement demand, considering factors such as total student population, backlogs, and credits. The model strives to forecast a student's placement likelihood, incorporating academic metrics, technical skills, and recruitment criteria.

Leveraging Predictive Analytics is paramount in refining placement predictions and shaping the development of students' skills. The profound influence of placement outcomes on a university's reputation underscores the need for specialized departments and preparatory classes. Within the realm of education, Machine Learning plays a pivotal role in forecasting placement trends through the analysis of historical data, providing invaluable insights for strategic planning. Notably, academic performance, especially in core subjects, emerges as a critical determinant in student placement, compelling institutions to prioritize both placement initiatives and academic records to bolster their standing and attract admissions.

2.RELATED WORK

This segment conducts a review of relevant literature, exploring research papers centered on campus placement utilizing supervised machine learning techniques.

[1] Student Placement Prediction Using Supervised Machine Learning:

Authors: M. Siva Surya, Dr. M. Sathish Kumar, Dr. D. Gandhimathi.

Description: This study explores the vital domain of student placement in academic institutions, emphasizing its impact on university admission and reputation. With a focus on improving placement services, the research utilizes recent pupil data and integrates a predictive algorithm to forecast current placement probabilities. Developed with the collaboration of the institution responsible for placement prediction and rigorous preprocessing, the proposed model showcases exceptional accuracy, surpassing conventional classification algorithms.

[2] Anticipating Placement Status of Engineering Students using Machine Learning based Prediction Models - A Case Study of Computer Science Stream:

Authors: Prof. Kajal Rai, Prof. Pawan Kumar, Prof. Sanjay Sood.

Description: This research addresses the crucial aspect of student placements in academic institutions, focusing on the

quality of placements, a key parameter for university rankings. The study employs machine learning techniques to predict the placement status of upcoming batches in the field of computer science engineering. Logistic Regression and Decision Tree algorithms exhibit the best predictive performance, identifying factors such as 'Xth' and 'XIIth' marks, 'Current CGPA,' and 'Present Attendance' as significant influencers on placement outcomes. The research emphasizes the importance of using intelligent computing to anticipate placement trends, providing valuable insights for enhancing overall placement quality.

[3] Placement Prediction System using Machine Learning:
Authors: Gurivireddy Mary Spandana, L Pallavi (Associate Professor).
Description: This study aims to improve the student placement system in academic institutions by introducing a predictive model. Utilizing machine learning algorithms such as Logistic Regression, Random Forest, KNN, and SVM, the research focuses on predicting student placement outcomes based on various parameters. The goal is to enhance overall system accuracy, and the study compares the performance of these algorithms on a common database. The results indicate that KNN achieves the highest accuracy at 75.90%, establishing itself as the most suitable algorithm for the given dataset. This predictive system holds significance for both student career development and the reputation of educational institutions.

[3] College Kart and k-NN Algorithm based Placement Prediction:
Authors: Nitesh Kumar Sharma, Aniket Kumar Singh, Shubham Salvi.
Description: This study focuses on student placement, a vital factor influencing educational institution admissions and credibility. The research introduces a system that utilizes the K-Nearest Neighbors (k-NN) supervised machine learning algorithm, achieving an 87% accuracy in predicting students' placement in companies. Alongside predictions, the system provides recommendations to improve placement prospects and includes a portal for students to buy, sell, and exchange academic materials. It also enables students to showcase skills, form project teams, and manage excess study materials. The research emphasizes the growing significance of placement in higher education, explores various machine learning techniques, and presents a block diagram illustrating the system's functionality. The k-NN algorithm and the system's performance are discussed, concluding with insights into benefits, limitations, and future enhancements like incorporating predicted online courses and coding assessments for comprehensive student evaluation.

[5] Campus Placement Prediction System Using Supervised Machine Learning Techniques:
Authors: Navuluri Divya, Sravya Namburu, Rajalakshmi Raja.
Description: This research focuses on student placement, vital for educational institutions' reputation and admissions. The study aims to predict current students' placement chances using data from previous years. A proposed algorithmic model, utilizing institution-specific data and preprocessing methods, is compared to traditional algorithms like Decision Tree and Random Forest. Results reveal the proposed algorithm's superior performance, achieving 75.90% accuracy. This research significantly contributes to improving placement predictions, benefiting both students and institutions striving for enhanced placement success.

3.METHODOLOGY

Initially, the project objectives are formulated, and data is gathered by the college. Subsequently, the collected data undergoes preparation, involving the segmentation into training and testing datasets. This process includes the segregation of data into dependent and independent variables, followed by the identification and removal of outliers.

The procedural steps are outlined as follows:

1. Training the Model:

The model undergoes training using the prepared training data.

2. Fine-Tuning the Model:

The model is refined, and the training data is further categorized based on the specified parameters.

3. Processing Testing Dataset:

Following model training, the previously stored testing dataset is processed through the model for prediction.

4. Prediction Confirmation:

If the predicted values from the testing dataset align with the expected outcomes, these values are confirmed as the final predictions.

5. Handling Unsatisfactory Predictions:

In cases where the predicted values are deemed unsatisfactory or false, indicating a lack of alignment with expectations, the values are returned for additional fine-tuning of the model.

This iterative process ensures continuous enhancement of the model's predictive capabilities, with adjustments made based on the outcomes of the testing dataset.

4.PROPOSED SYSTEM

Employed Algorithms:

- **Logistic Regression**
- **Random Forest**
- **K-Nearest Neighbors (KNN)**
- **XGBoost**

Logistic Regression (Binary Classification):

Logistic regression, a commonly utilized statistical technique, is applied to forecast binary outcomes, such as the probability of an individual securing a job post-training. The methodology entails collecting information on individuals who have finished the training program, encompassing details like educational background, professional experience, and performance during the training initiative. Predictions are generated by inputting the predictor values for a new individual, with the model estimating the probability of job placement. If the probability surpasses a designated threshold (e.g., 50%), the system anticipates successful placement; otherwise, it predicts non-placement.

Random Forest:

Random Forest, a versatile algorithm ideal for multiclass classification, is harnessed to predict students' placement ranges.

Data Preparation: Thorough data on academic performance, skills, and relevant parameters is amassed.

Feature Selection: Crucial features correlated with the target variable (placement range) are pinpointed, employing techniques like correlation analysis.

Model Training: An ensemble of decision trees is iteratively crafted to minimize the loss function.

Hyperparameter Tuning: Fine-tuning of hyperparameters, encompassing the learning rate, number of trees, and maximum depth, is undertaken.

Model Evaluation: Comprehensive assessment of model performance is conducted, incorporating metrics such as accuracy, precision, recall, and F1-score.

K-Nearest Neighbors (KNN):

KNN is strategically applied in multiclass classification for forecasting student placement.

Data Processing: Collection and preprocessing of data, encompassing addressing missing values, outliers, and encoding categorical variables, is meticulously executed.

Feature Importance: Identification of pivotal features is accomplished using techniques such as mutual information.

Train-Test Split: Strategic division of data into training and testing sets (e.g., 80-20 split) is implemented.

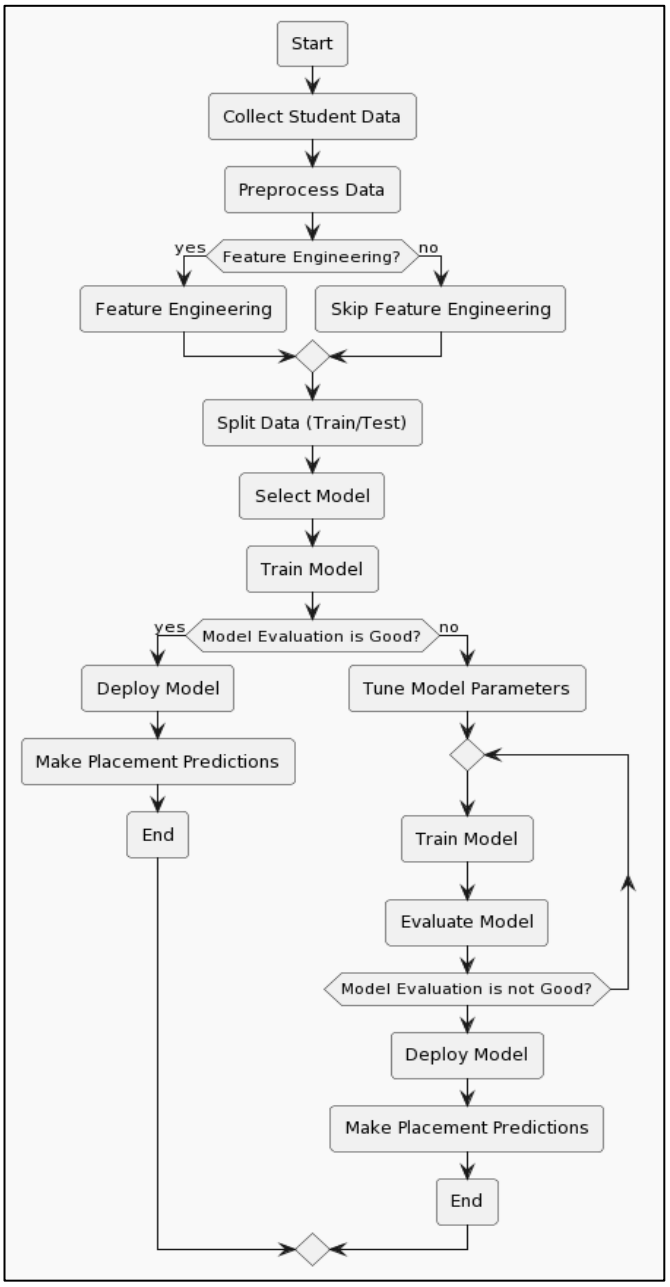
Model Training: Training of KNN classifiers on the provided dataset is systematically carried out.

Hyperparameter Tuning: Optimization of hyperparameters specific to the KNN algorithm is diligently performed.

Evaluation: Rigorous assessment of performance metrics, including precision, recall, F1-score, and accuracy, is executed.

Our future scope involves harnessing the advanced capabilities of XGBoost, also referred to as eXtreme Gradient Boosting, to substantially enhance predictive accuracy. By leveraging the strengths of this sophisticated machine learning algorithm, we anticipate achieving more precise and reliable predictions in our upcoming analytical endeavors.

5.SYSTEM ARCHITECTURE



6. DISCUSSION

1. Integration of Data Analytics and Machine Learning: Combining data analytics and machine learning improves predicting student placements. Analytics mines placement data for patterns, while machine learning creates adaptable algorithms based on this data.

2. Holistic Factors Approach: Successful prediction models consider diverse factors like academic performance, soft skills, adaptability, and practical experience for better employability forecasts.

3. Soft Skills as Predictors: Recognizing soft skills as crucial predictors reflects a shift in education. Including communication, teamwork, adaptability, and problem-solving in predictive models adds depth to analysis.

4. Challenges in Predictive Analysis: Challenges like data quality and model interpretability must be addressed for accurate predictions. Ensuring reliable input data and transparent models is crucial, especially in educational settings.

5. Ethical Considerations: Addressing biases in algorithms, ensuring data privacy, and maintaining transparency are vital ethical considerations. Balancing predictive analysis benefits with ethical concerns is essential for responsible implementation.

6. Case Studies and Success Stories: Examining real cases offers practical insights for effective predictive models. Learning from successful implementations guides future strategies and methodologies.

7. Emerging Trends and Recommendations: Identifying new trends like advanced AI, natural language processing, and predictive analytics in remote learning shapes future directions for placement predictions.

8. Institutional Adaptation and Implementation: Educational institutions need to invest in resources, training, and collaboration with industry partners to effectively adopt and maintain predictive models aligned with industry demands.

6. CONCLUSION AND FUTURE SCOPE

In conclusion, our research underscores the pivotal role of predictive analytics in optimizing student campus placement. The strategic deployment of advanced machine learning algorithms, including Logistic Regression, Random Forest, and K-Nearest Neighbors, has yielded a substantial improvement in accuracy when compared to conventional methodologies.

The proposed system, designed for comprehensive student evaluation, integrates meticulous data preprocessing and iterative model refinement. Looking ahead, the integration of XGBoost represents a promising avenue for further enhancing predictive accuracy. Our study significantly contributes to the progression of placement predictions, providing tangible benefits to both students and institutions.

Future research directions involve harnessing XGBoost's advanced capabilities for precise predictions and exploring its broader applications in analytics and decision-making processes within educational contexts.

7.REFERENCES

[1] M. Siva Surya, Dr. M. Sathish Kumar, Dr. D. Gandhimathi "Student Placement Prediction Using Supervised Machine Learning"
DOI: 10.1109/ICACITE53722.2022.9823648.

[2] Gurivireddy Mary Spandana, L Pallavi "Placement Prediction System using Machine Learning"
DOI: 10.1109/ICECAA58104.2023.10212409.

[3] Navuluri Divya, Sravya Namburu, Rajalakshmi Raja "Student Placement Analysis using Machine Learning" DOI: 10.1109/ICCES57224.2023.10192633.

[4] Jyoti Khandelwal Gaurav Pareek, Ratul Dey, Shilpa Pareek "The Study of Machine Learning Classification Algorithm for Student Placement Prediction"
DOI: 10.1109/IEMECON56962.2023.10092294.

[5] K. Rai, Predicting Covid-19 based on symptoms using machine learning techniques implemented in Python, AIP Conference Proceedings, vol. 2555, pp. 050002-1–050002-10, October 2022.

[6] Manoj Manike, Priyanshu Singh, Purna Sai Madala, Steve Abraham Varghese, Saleti Sumalatha "Student Placement Chance Prediction Model using Machine Learning Techniques" DOI: 10.1109/CICT53865.2020.9672372.

[7] B. Singh, S. K. Henge, A. Sharma, C. Menaka, P. Kumar, S. K. Mandal, and B. Debtera, ML-Based Interconnected Affecting Factors with Supporting Matrices for Assessment of Risk in Stock Market, Wireless Communications and Mobile Computing, vol. 2022, 15 pages, Dec. 2022.

[8] Nitesh Kumar Sharma, Aniket Kumar Singh, Shubham Salvi "College Kart and k-NN Algorithm based Placement Prediction"
DOI: 10.1109/ICESC51422.2021.9532987