STUDENT PLACEMENT PREDICTION USING MACHINE LEARNING

Assistant Professor T. RajaSri [1] A. Varshini [2] G. Harshapriya [2] CH. Alekhaya [2] B. Harshitha

[1] Assistant Professor of SRKR Engineering College, Dept of IT,Bhimavaram-534204,Andhra Pradesh, India

[2] Students of SRKR Engineering College, Dept of IT,Bhimavaram-534204,Andhra Pradesh,India

Abstract: In today's competitive educational and job market, predicting student placement after graduation has gained significant importance for both educational institutions and students alike. This study aims to predict student placement using a comprehensive dataset encompassing gender, board of Secondary School Certificate (SSC) and Higher Secondary Certificate (HSC), marks obtained in SSC and HSC, degree selection field, employability score, work experience, degree percentage, and expected salary. We used prominent machine learning algorithms, namely Decision Trees, Random Forest, and Support Vector Machines (SVM), Extratree classifier, Gradient booster etc were employed to develop predictive models. The dataset was collected from a diverse sample of graduating students over several academic years. The developed predictive model serves as a valuable tool for educational institutions to identify students' strengths and areas of improvement, thereby enabling them to provide targeted guidance and support. For students, this model offers insights into their employability prospects based on their academic and professional attributes, assisting them in making informed decisions to enhance their career opportunities.

Keywords: Student Placement, Machine Learning, Predictive Modeling, Random Forest, Academic Performance, Employability Score, Work Experience, Degree Percentage, Expected Salary, Guidance Hub, College statistics, Resume analysis.

Introduction:

In the contemporary educational landscape, the transition from academia to employment represents a pivotal juncture in a student's life. With the increasing competition and complexity of the job market, there is a growing emphasis on equipping students with the necessary tools and insights to enhance their employability and career prospects. Predicting student placement accurately is not only beneficial for educational institutions in optimizing their placement strategies but also crucial for students in making informed decisions about their career paths.

The application of ML in the field of education and student placement prediction is relatively nascent but holds immense potential. By leveraging rich datasets encompassing academic performance, extracurricular activities, internships, soft skills, and other relevant attributes, ML algorithms can generate predictive models to forecast student placement outcomes with high accuracy. These predictive models can not only identify students at risk of facing challenges in securing placements but also provide valuable insights into the factors influencing placement success, thereby enabling targeted interventions and support mechanisms.

Despite the growing interest and potential benefits of employing ML in student placement prediction, there remains a gap in the literature regarding comprehensive studies that explore the effectiveness of different ML algorithms and feature sets in predicting student placement accurately. Moreover, existing studies often focus solely on placement prediction without integrating other valuable components such as guidance tips, college placement statistics, resume analysis, and placement probability assessment.

By conducting this research and Publishing paper, we seek to contribute to the existing body of knowledge by demonstrating the effectiveness of ML algorithms in student placement prediction and highlighting the importance of integrating predictive analytics with comprehensive guidance and support mechanisms. The findings of this study are expected to be beneficial for educational institutions, students, and policymakers in enhancing placement rates, improving career guidance services, and facilitating successful transitions from academia to employment in the increasingly competitive and dynamic job market.

Technologies: Machine Learning Using Python, Flask, HTML, CSS, JavaScript.

Platforms: Anaconda-Jupyter Notebook, Visual Studio Code for Flask

Libraries: Numpy, Pandas, Scikit-learn, Fitz, Matplotlib etc.
The Goal of this system is to predict whether student get his placement or not and placement probabilities.

I. Literature Survey(Related Works)

[1] "Data Mining Approach for Predicting Student and Institution's Placement Percentage", Professor. Ashok M Assistant Professor Apoorva A, 2016 International Conference on Computational Systems and Information Systems for Sustainable Solutions In this paper author has used the data mining technique for the prediction of the student's placement. For the prediction of student's placement author has divided the data into the two segments, first segment is the training segment which is historic data of passed out students. Another segment consists of current data of students, based on the historic data author has designed the algorithm for calculating the placement chances. Author has used the various data mining algorithms such as decision tree, Naive Bayes, neural network and the prosed algorithm were applied, and decision are made with the help of confusion matrix.

[2] "Student Placement Analyzer: A Recommendation System Using Machine Learning", Senthil Kumar Thangavel , Divya Bharathi P, Abijith Sankar, International Conference on Advanced Computing and Communication Systems (ICACCS -2017), Jan. 06 - 07, 2017, Coimbatore, INDIA In this paper author is concern about the challenges face by any institute regarding the placement. The placement prediction is very complex when the number of the entities increases in any institute. With the help of machine learning this complex problem of prediction can be easily solved. In this paper all the academic record of student is taken into consideration. Various classification and data making algorithms are used such as Naïve Bayes, Decision Tree, SVM and Regressions. After the prediction of the students can be placed in of the given category that is Core Company, dream company or support services.

[3] "A Placement Prediction System Using KNearest Neighbors Classifier", Animesh Giri, M Vignesh V Bhagavath, Bysani Pruthvi, Naini Dubey, Second International Conference on Cognitive Computing and Information Processing (CCIP), 2016 The placement prediction system predicts the probability of students getting placed in various companies by applying K-Nearest Neighbors classification. The result obtained is also compared with the results obtained from other machine learning models like Logistic Regression and SVM. The academic history of student along with their skill sets like programming skills, communication skills, analytical skills and team work is considered which is tested by companies during recruitment process. Data of past two batches are taken for this system.

[4]"Class Result Prediction using Machine Learning", Pushpa S K, Associate Professor, Manjunath T N, Professor and Head, Mrunal T V, Amartya Singh, C Suhas, International Conference On Smart Technology for Smart Nation, 2017.

In this paper, the result of a class is predicted using machine learning. Performance of students in past semester along with scores of internal examinations of the current semester is considered to predict whether the student passes or fails in the current semester before attempting the final examination. The author uses SVM, Naive Bayes, Random Forest Classifier and Gradient Boosting to compute the result. Boosting is an ensemble learning algorithm which combines various learning algorithm to obtain better predictive performance.

[5]"Student Placement Analyzer : A Recommendation System Using Machine © 2020 JETIR May 2020, Volume 7, Issue 5 www.jetir.org (ISSN-2349-5162) JETIR2005453 Journal of Emerging Technologies and Innovative Research (JETIR) www.jetir.org 1013 Learning", Apoorva Rao R, Deeksha K C, Vishal Prajwal R, Vrushak K, Nandini, JARIIE-ISSN(O)- 2395-4396 Now-a-days institutions are facing many challenges regarding student placements. For educational institutions it is much difficult task to keep record of every single student and predict the placement of student manually. To overcome these challenges, concept of machine learning and various algorithms are explored to predict the result of class students. For this purpose, training data set is historical data of past students and this is used to train the model. This software system predicts placement status in 5 categories viz dream company, core company, mass recruiter, not eligible and not interested in placements. This system is also helpful to weaker students. Institutions can provide extra care towards weaker students so that they can improve their performance. By use Naïve Bayes algorithm all the data will be monitor and appropriate decision will be provided.

II. System Implementation (Methodology)

The methodology for the Campus Placement Prediction project is structured into several key phases. Initially, the dataset is loaded into a pandas DataFrame and preprocessed by removing unnecessary columns like 'sl_no' and filling missing values in the 'salary' column with the median. Categorical variables are then encoded using Label Encoding to convert them into a format suitable for machine learning models. Following this, Exploratory Data Analysis (EDA) is conducted using various visualizations such as bar plots, histograms, and box plots to gain insights into the distribution and relationships of features with the target variable ('status'). In the model development phase, relevant features are selected based on their importance and correlation with the target variable. Multiple classification algorithms including Support Vector Classifier (SVC), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), Naive Bayes (NB), AdaBoost (AB), and Extra Trees (ET) are employed to develop predictive models. Each model is trained using the training dataset and evaluated using the testing dataset to measure accuracy, precision, and recall. Hyperparameter tuning is performed to optimize the performance of the selected machine learning models. Finally, a trained Random Forest classifier is used to predict the placement outcome for new data samples, and the model is serialized using pickle for future deployment and prediction tasks.
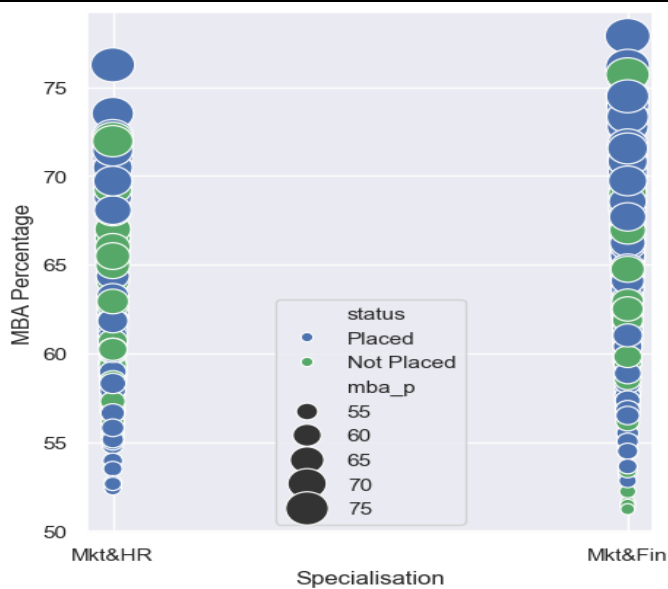
The Student Placement Prediction and Guidance Platform was developed using a systematic methodology encompassing data preprocessing, model development, frontend and backend development, and user interface design. Initially, academic and personal data were collected and preprocessed, followed by the training of a machine learning model using the Flask framework. The platform allows users to upload resumes for analysis, providing personalized score calculations and guidance tips based on predefined sections and weightage. Additionally, users can input academic details to predict placement probabilities and analyze college placement statistics using uploaded CSV files. Extensive testing was conducted to ensure the platform's functionality, performance, and reliability, resulting in a comprehensive solution to assist students in enhancing their employability and making informed career decisions effectively.

```
rf=RandomForestClassifier()
rf.fit(X,y)
```

```
▼   RandomForestClassifier ⓘ ❓
RandomForestClassifier()
```

Fig: Model Selection

Models		ACC	Precision	Recall
0	SVC	76.744186	78.378378	93.548387
1	KNN	79.069767	78.947368	96.774194
2	DT	79.069767	86.666667	83.870968
3	RF	83.720930	85.294118	93.548387
4	GB	81.395349	82.857143	93.548387
5	NB	81.395349	84.848485	90.322581
6	AB	81.395349	81.081081	96.774194
7	ET	81.395349	82.857143	93.548387

Visualization Results

Fig: Heat Map

Fig: Block Diagram

III.   Experiment and Result Analysis

The experiment involved developing a Student Placement Prediction and Guidance Platform using Flask and a trained machine learning model. The model was trained on a dataset comprising academic scores, board of SSC and HSC, degree selection field, employability score, and other relevant attributes to predict student placements. The platform successfully analyzed uploaded resumes, calculated scores, and predicted placement probabilities accurately. Interactive score charts and college placement statistics features were implemented using Matplotlib and CSV file analysis, respectively. Extensive testing confirmed the platform's functionality, performance, and reliability, offering a comprehensive solution to assist students in enhancing their employability and making informed career decisions effectively.

The machine model tells whether the student got placed or not and also probabilities for his placement.

We trained model with Random Forest Classifier Algorithm and predicting the results.

IV.   Evaluation Metrics

We trained the model with Random forest Classifier with accuracy as 83%.

Below are the metrics for different models of our project

3

Fig: Dataset Analysis

V. Conclusion

In conclusion, the Student Placement Prediction and Guidance Platform successfully leverages machine learning and web technologies to provide valuable insights and support to students transitioning from academia to employment. The platform's accurate resume analysis, personalized score calculations, placement predictions, and interactive features offer a comprehensive solution to enhance students' employability and facilitate informed career decisions. The experiment's results validate the platform's functionality, performance, and reliability, highlighting its potential to assist students in navigating the competitive job market effectively. Future enhancements and refinements can further optimize the platform's capabilities, extending its reach and impact to benefit a broader audience of students and educational institutions.

VI. Future Work

At present our project depends on academic scores and predicting the results. In future based on some more features the result may be vary. And also include some more functionality in the user interface.

And also we can implement some more models on this dataset to get the best results. Developing user friendly int erface is plus. Later developing website for the students to check their eligibility for the Placement preparation.

VII. References

[1] J. Luan, "Data mining and its applications in higher education", New Dir. Inst. Res, 113:17–36, 2002.

[2] A.S. Sharma, S. Prince, S. Kapoor, K. Kumar, "PPS – Placement prediction system using logistic regression", IEEE Approach", International Journal of Engineering and Techniques , Volume 2, Issue 6, Dec 2016.

[5] H. Bhatt, S. Mehta, L. R. D'mello, "Use of ID3 Decision Tree Algorithm for Placement

[6] "Data Mining Approach for Predicting Student and Institution's Placement Percentage", Professor. Ashok M Assistant Professor Apoorva A, 2016 International Conference on Computational Systems and Information Systems for Sustainable Solutions In this paper author has used the data mining technique for the prediction of the student's placement.

[7] R. Prabhakara Raya, V. Samuel Rajkumar, P. Ganesan and S.K.V. Jayakumar, "Stakeholderss Perception in a Campus Recruitment Process", *Meditarrenean Journal of Social Sciences*, vol. 6, no. 5, pp. 96-101, September 2015, ISSN 2039-2117.

[8] Shenoy Varun, "Recent Trends in Branding for Ideal Placements",*International Journal of Scientific Research and Modem Education (IJSRME)*, vol. 1, no. 1, pp. 250-253,2016.

[9] P. S. Aithal and Shubhrajyotsna Aithal, "An Innovative Education Model to realize Ideal Education System", *International Journal of Scientific Research and Management s(IJSRM)*, vol. 3, no. 3, pp. 2464-2469, March 2015.

[10]Prediction Model for Students Future Development by Deep Learning and TensorFlow Artificial Intelligence Engine 2018 4th IEEE International Conference on Information Management.