PROJECT REPORTON

“Campus Placements Prediction”

Submitted by
TEAM: Code Compilers
(M Keerthana
T Varshini
Y Sathwika)

Table of Contents

1. INTRODUCTION

Overview:

In the rapidly evolving landscape of education and industry, the transition from academia to the professional world is a critical phase for students. Campus placements stand as a pivotal bridge that connects the theoretical knowledge gained in classrooms to the practical challenges posed by the corporate sphere. However, this process is intricate and multifaceted, impacted by a myriad of factors that influence the successful placement of students into desired job roles.

This is where our application “CAMPUS PLACEMENT PREDICTION” comes in.The challenge of predicting whether a candidate will be placed or not leverages the power of data-driven insights to provide informed guidance to both students and recruiters. By analyzing historical campus placement data, incorporating various attributes, and employing advanced algorithms, we aim to develop predictive models that can estimate the probability of a candidate's placement success.

The approach involves a data-centric journey, starting with data collection, preprocessing, and feature engineering. We will delve into various machine learning algorithms, including classification techniques, to create predictive models. These models will be trained on historical placement data, learning the intricate patterns that contribute to successful placements.

Purpose*:*

The purpose of identifying patterns and trends in campus placement data using machine learning is to harness the power of data-driven insights to enhance the effectiveness of the placement process. By analyzing historical placement data, we aim to uncover hidden correlations and influential factors that impact successful placements.

This endeavor enables educational institutions to tailor their curricula and support systems to better align with industry demands, thus improving students' employability. Recruiters, on the other hand, can make more informed decisions by predicting which candidates are likely to succeed in specific roles, optimizing their selection process and resource allocation. Ultimately, this purpose strives to bridge the gap between academia and industry, empowering both students and employers through data-driven decision-making and fostering a more strategic and student-centric approach to campus placements.

The purpose extends to recruiters, enabling them to streamline their efforts and resources toward candidates who exhibit a higher likelihood of flourishing in specific roles. Moreover, this initiative seeks to cultivate a culture of continuous improvement, where data-driven insights prompt refinements in strategies, enhancing the overall placement process. Ultimately, this purpose extends beyond data analysis – it aims to drive a profound transformation that harmonizes education, recruitment, and career growth through the lens of insightful patterns.

2. LITERATURE SURVEY

a. Existing Problem:

In the context of campus placement prediction, there are a number of existing problems that need to be addressed. Some of these problems include:

1) Lack of data: As mentioned earlier, there is often not enough data available to train a machine learning model to accurately predict campus placement. This is because the number of students who get placed is relatively small, and the data about their academic performance, skills, and other factors is often incomplete or inaccurate.

2) Non-linear relationships: The relationship between a student's academic performance and their chances of getting placed is often non-linear. This means that a small change in academic performance can have a disproportionately large impact on the placement chances. This can make it difficult to predict placement accurately using machine learning models.

3)Uncertainty: There is always some uncertainty involved in predicting campus placement. This is because there are many factors that can affect a student's chances of getting placed, and it is impossible to predict all of these factors with certainty.

4) Bias: Machine learning models can be biased, which means that they may not accurately predict the placement chances of certain groups of students. This bias can be caused by the data that is used to train the model, or by the way that the model is designed.

b. Proposed Solution:

Predicting campus placements involves analyzing various factors that can influence a student's likelihood of getting placed. Here's a proposed solution using a machine learning approach:

Data Collection: Gather data from previous campus placements, including student profiles (academic records, skills, etc.), company profiles (industry, job roles offered, etc.), and placement outcomes (placed or not).

Data Preprocessing: Clean and pre process the data. Handle missing values, normalize numerical features, and encode categorical variables

Feature Selection/Engineering: Identify relevant features that can impact placement outcomes.

Model Selection: Choose appropriate machine learning algorithms. Given that this is a binary classification problem (placed or not), consider algorithms like: Logistic Regression, Random Forest, Support Vector Machines, Gradient Boosting

Model Training: Split the data into training and testing sets. Train the selected models on the training data.

Model Evaluation: Evaluate the model's performance on the testing data using metrics like accuracy, precision, recall, F1-score, and ROC-AUC.

Hyper parameter Tuning: Optimize model hyper-parameters using techniques like grid search or random search to improve performance.

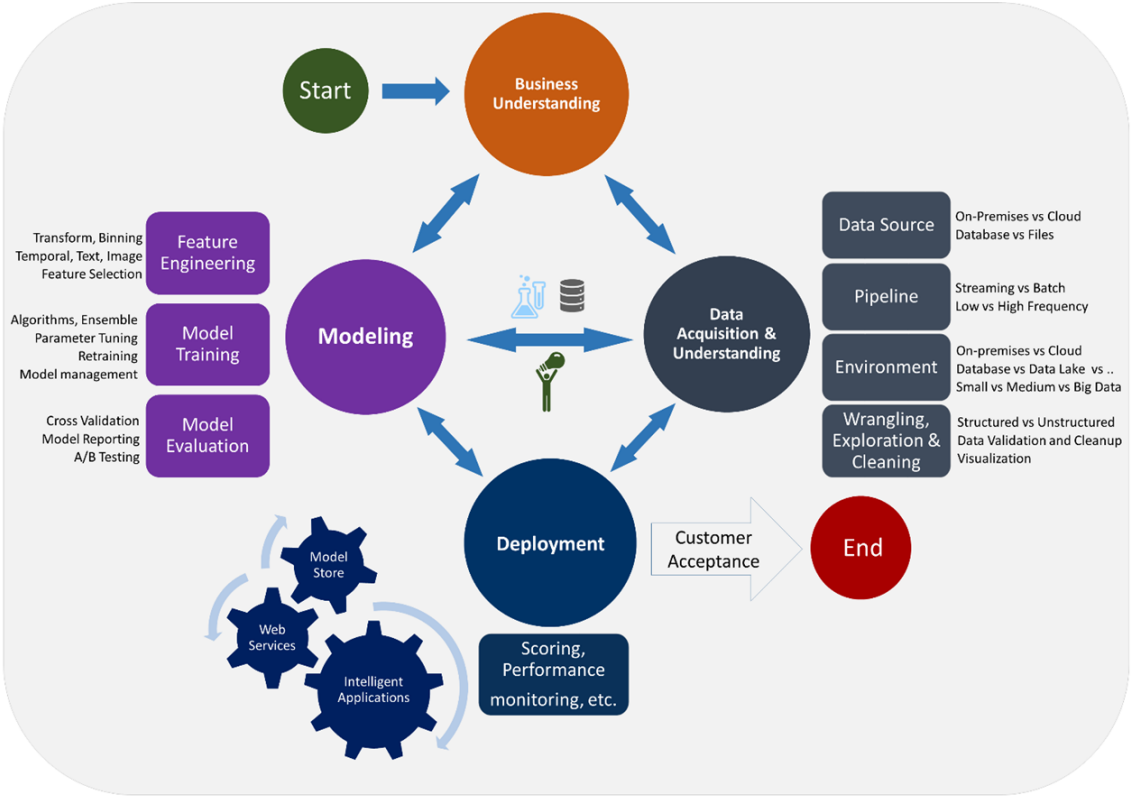
Interpretability: Depending on the chosen algorithm, interpret the model's decisions to understand the importance of different features.

Deployment: Once satisfied with the model's performance, deploy it to make real-time predictions. This could be through a web application or API.

Continuous Learning: Collect new placement data and periodically retrain the model to keep it up-to-date and accurate.

3. THEORITICAL ANALYSIS:

3.1 BLOCK DIAGRAM:



3.2 HARDWARE/SOFTWARE REQUIREMENTS:

HARDWARE: This web application doesn't need any extra hardware components in addition to the default version. The basic computer components are enough to make a good use of this application. This is a software project.

SOFTWARE: We talk about the software components in detail. We categorize them based other usage and we describe the stack we used.
FRONT END : Front end is nothing but designing user interface of the application.TECH **STACK** : Html, CSS, JavaScript

BACK END: Back end is designing the overall login behind the workflow of the project. Itis a place where all the use case functionalities were written
TECH STACK : Python Flask

DATABASE: Database is a place where we store the required data in the tabular format by forming some schemas based on Relational Database Management System(RDBMS).Whenever the storage platform is a cloud platform. We call it as a cloud database. We are using cloud database in this project
TECH STACK: IBM DB2

DEPLOYMENT : We deployed our ML model using flask. It is a popular approach to make the model accessible through a web API.
TECH STACK: Flask

4. EXPERIMENTAL INVESTIGATIONS:

Collect data: We collected data about the students. This data can include their academic performance, skills, personality traits, extracurricular activities, and other factors that may be relevant to their chances of getting placed.

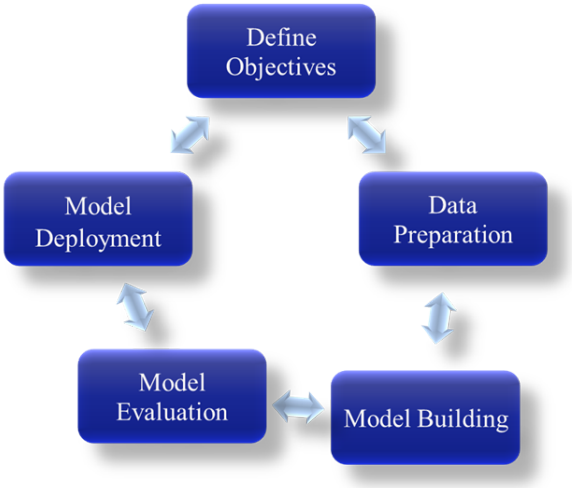
Choose a machine learning algorithm: There are many different machine learning algorithms that can be used for campus placement prediction. The choice of algorithm will depend on the data that is available and the specific factors that are being considered.

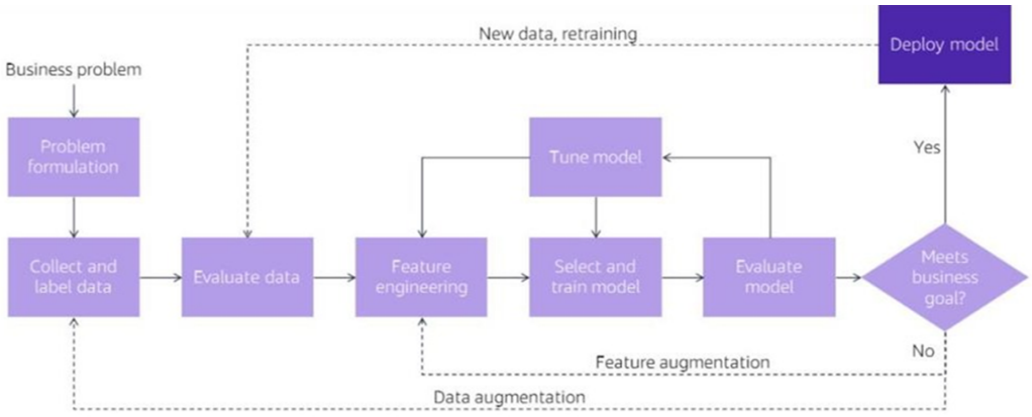
Train the model: The machine learning algorithm is then trained on the data that has been collected. This training process involves the algorithm learning to identify patterns in the data that are associated with placement.

Test the model: Once the model has been trained, it is important to test it on a separate dataset of data that was not used to train the model. This will help to ensure that the model is not overfitting to the training data and that it is able to generalize to new data.

Evaluate the model: he performance of the model can be evaluated using a variety of metrics, such as accuracy, precision, and recall. These metrics measure how well the model is able to predict the placement of students.

5. FLOWCHART:

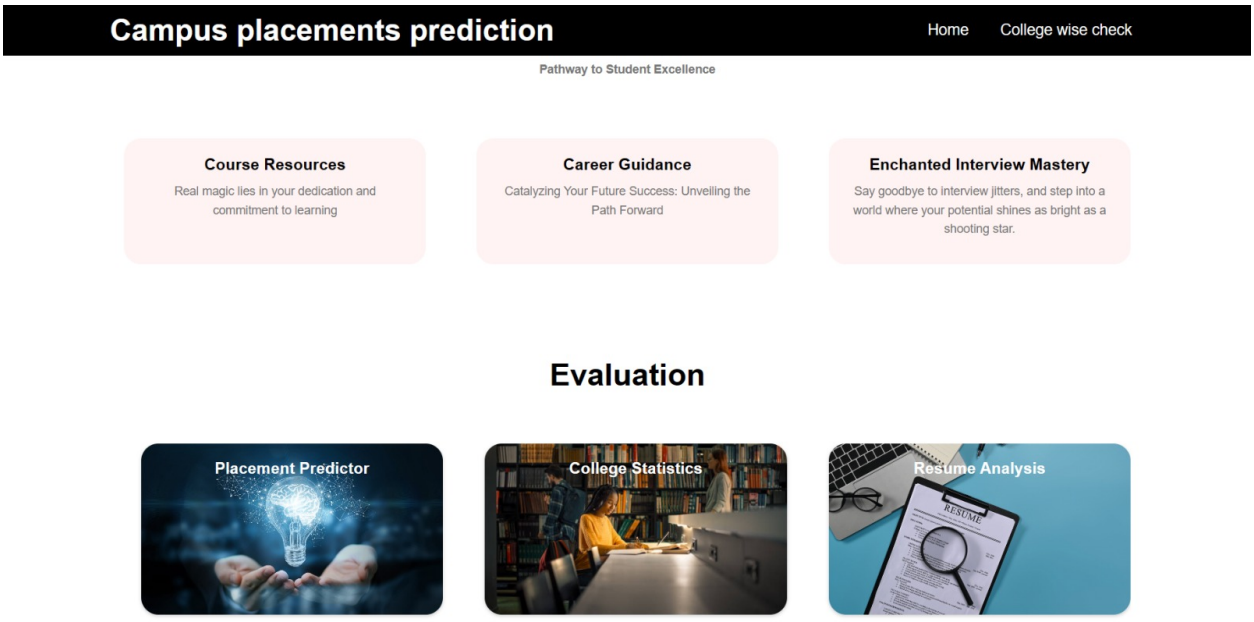

Define Objectives
Data Preparation
Model Building
Model Evaluation
Model Deployment


Business problem
Problem formulation
Collect and label data
Evaluate data
Feature engineering
Select and train model
Tune model
Evaluate model
Meets business goal?
Yes
No
Deploy model
New data, retraining
Feature augmentation
Data augmentation

6. RESULT

The outcomes of campus placements shown along with screenshots. Much more pages are present such as Home, Registration form,career guidance, interview tips, course resources and much more.

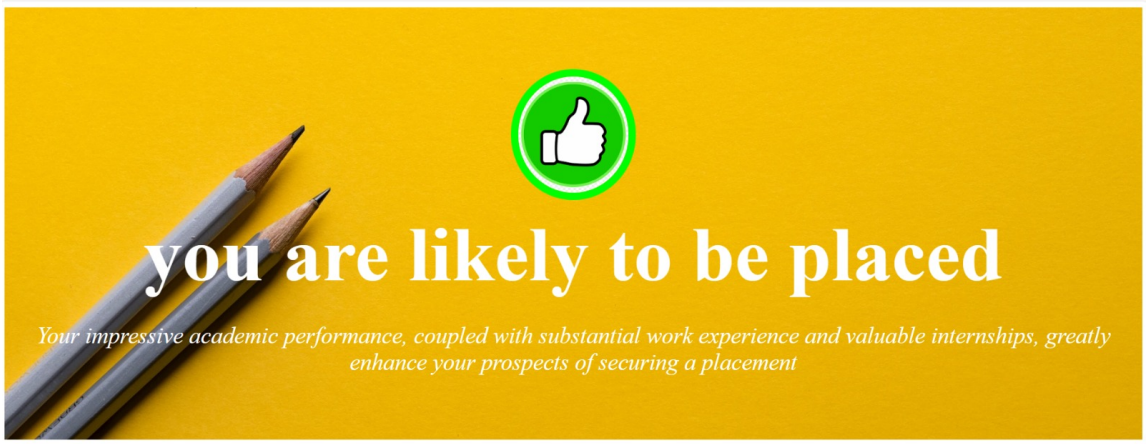
Home Page



The main outcomes of our project are Placement Predictor, College Statistics and Resume Analysis

Placement Predictor:

By considering all the factors which are going to influence placement success if the person has more chances to get the placement we will get the following output

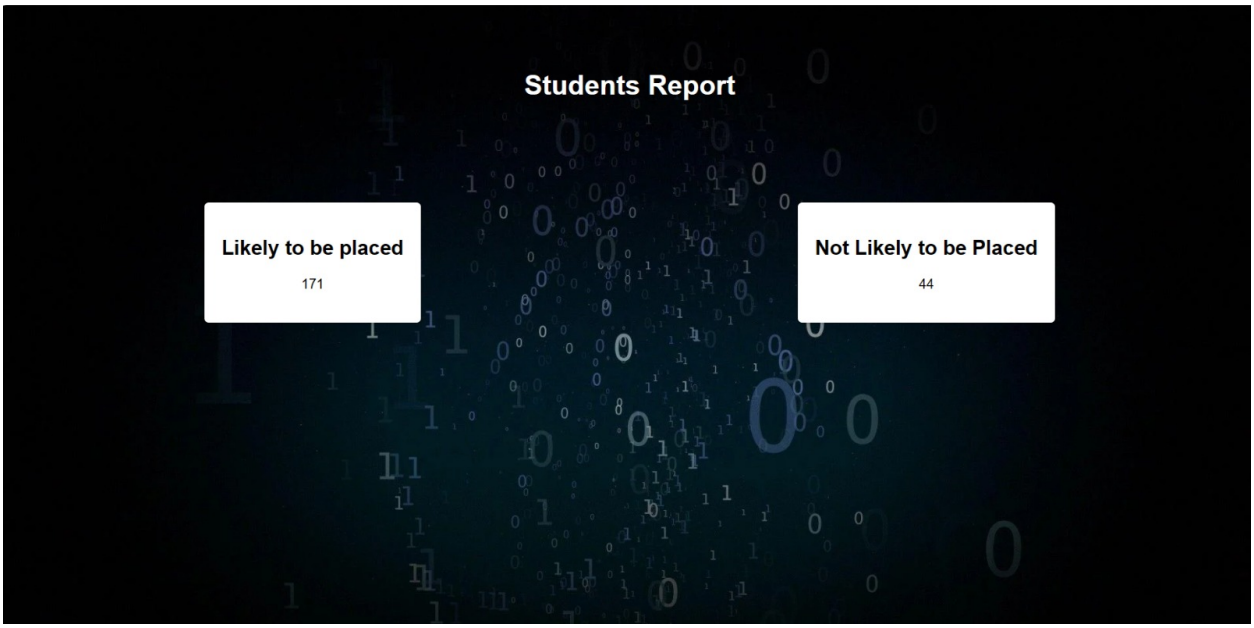


By considering all the factors which are going to influence placement success if the person has less or no chances to get the placement we will get the following output

College Statistics:

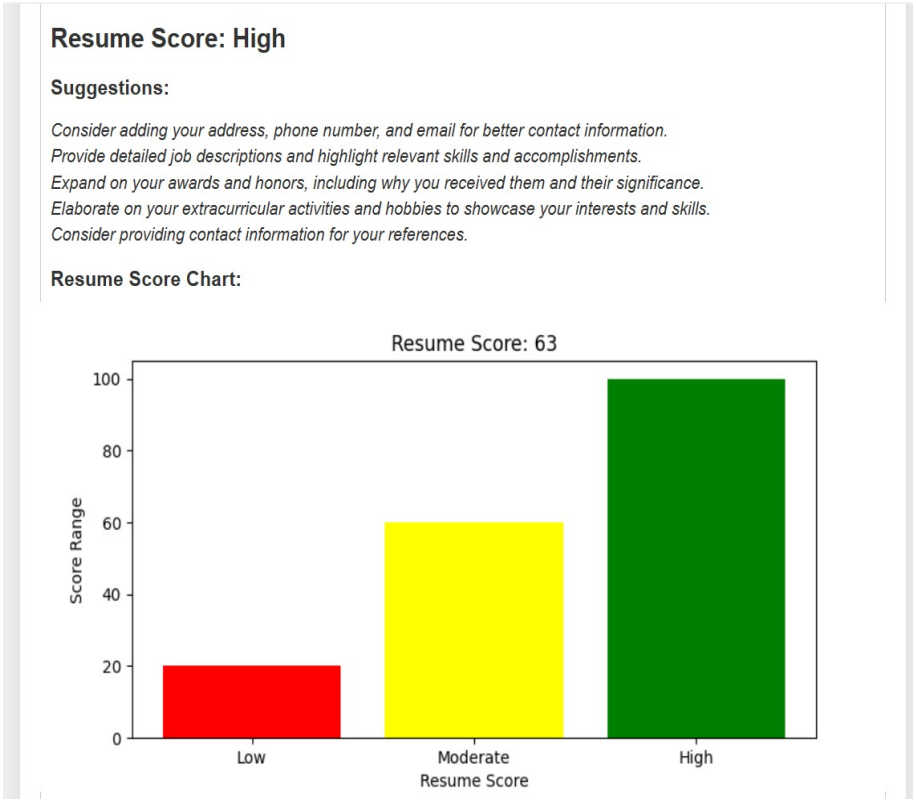
After uploading students information of college according to our dataset it will display the number of students likely to be placed and number of students who have less chances to get placed. This will give a clear insight about student placement to the institution.

The output of this is shown below



Resume Analysis:

Resume analysis for students can be a valuable tool for several purposes, including career development, job search preparation, and academic planning.



7. ADVANTAGES AND DISADVANTAGES

ADVANTAGES:

USER FRIENDLY INTERFACE: Campus Placement Prediction is designed with a simple user interface to allow users to work with ease. With a simple glance at the software, users can understand how to use it. No formal training is required for the employees as the software can beinstalled and operated without technical knowledge.

SECURE: Campus Placement Prediction provides a safe and secure medium to save your data. The billing application uses cloud-based servers to store the information. Data safety is taken utmost care of by saving it on anonymous servers. Even when a user loses thedevice, the data can be restored on any new device by using the same login credentials.

LOW COST: Maintenance Campus Placement Prediction of doesn't cost you much. You just need to have a cloud database plan in IBM DB2 which is available at reasonable prices.

FAST: You can enter the details into this application with ease by simply selecting the options present over there without any effort. You can do it anytime anywhere even when you are sitting in your lawn having a cup of coffee

DISADVANTAGES:

CONFINED TO SINGLE CANDIDATE: This is the only disadvantage for our "Campus Placement Prediction ".But with somechanges in the schema, we can extend this application to Various Colleges.

8. APPLICATIONS:

The application of identifying patterns and trends in campus placement data using machine learning techniques has far-reaching implications across various domains:

Education Institutions: Institutions can tailor their curricula to match industry demands, ensuring that students graduate with skills that are highly sought after by employers.

Recruiters and Employers: Recruiters can focus their efforts on candidates with a higher likelihood of success, optimizing their recruitment process and saving time and resources.

Students: Students gain insights into their strengths and areas for improvement, helping them make informed decisions about career paths and skill development.

Placement Agencies: Placement agencies can optimize candidate-employer matching by using predictive models to identify the best-fit candidates for specific roles.

8.CONCLUSION:

Campus placement prediction is a complex task that can be improved by using machine learning algorithms.The accuracy of the model depends on the availability of data, the quality of the data, the features used, and the machine learning algorithm.It is important to consider the limitations of campus placement prediction, such as the availability of data and the non-linear relationships between the different factors.By considering these factors, we developed a campus placement prediction mode,Resume score predictor that is accurate and reliable.

10. FUTURE SCOPE:

We can extend the statistics application for i.e we can evaluate and give the list of students succeded in the placements of that particular college. This can be done by correct modification of schema and we can build it in the future.

11. BIBILIOGRAPHY :

a. REFERENCES:

https://smartinternz.com/
https://www.ibm.com/academic/

b. APPENDIX:

SOURCE CODE:
Github: https://github.com/smartinternz02/SBSPS-Challenge-10430-1691070545