CAMPUS PLACEMENT PREDICTION

High Level Design (HLD)

Contents

Document Version Control

Version	Date	Author	Comments
1.0	**24-Mar-24**	**Bhavika Pathak**	**Introduction & Architecture defined**

Abstract

The High-Level Design for the Machine Learning Campus Placement Prediction Project aims to revolutionize the recruitment process within educational institutions by leveraging advanced machine learning techniques. The project focuses on developing predictive models capable of accurately forecasting placement outcomes for students based on a variety of input parameters such as academic performance, skill sets, and demographic factors.

The design encompasses several key components, including data collection and preprocessing, feature engineering, model selection and training, and deployment. Historical placement data, along with supplementary information such as student profiles and recruiter feedback, is collected and meticulously preprocessed to ensure data quality and consistency. Feature engineering techniques are then applied to extract relevant features and enhance the predictive power of the models.

Multiple machine learning algorithms, including but not limited to regression, classification, and ensemble methods, are considered for model selection. These algorithms are trained on the preprocessed data using appropriate evaluation metrics to optimize performance and generalization capability. The selected models undergo rigorous testing and validation to assess their accuracy, robustness, and scalability.

Finally, the trained models are deployed into a production environment, where they are integrated into a user-friendly interface accessible to students, recruiters, and administrators. The interface provides intuitive visualization of placement predictions and offers actionable insights to aid decision-making. Continuous monitoring and refinement mechanisms are implemented to ensure the ongoing relevance and effectiveness of the predictive models.

Overall, the High-Level Design for the Machine Learning Campus Placement Prediction Project serves as a comprehensive roadmap for developing and deploying a sophisticated predictive system that has the potential to streamline the recruitment process, enhance student outcomes, and optimize resource allocation within educational institutions.

1. Introduction

1.1. Why High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the important details about this project. Through this HLD Document, I'm going to describe every small and big thing about this project.

2. General Description

2.1. Product Perspective

The purpose of credit card default prediction project is to develop predictive models that accurately identify customers at risk of defaulting on their credit card payments.

2.2 Problem statement

Financial threats are displaying a trend about the credit risk of commercial banks as the incredible improvement in the financial industry has arisen. In this way, one of the biggest threats faces by commercial banks is the risk prediction of credit clients. The goal is to predict the probability of student placement based on collage student history.

2.3 Proposed Solution

In the proposed solution, we used SVM machine learning model to classify the student placement based on collage student data history. Here, first we are performing Data preprocessing step, in which data transformation, handling missing values, feature transformation, feature selection, steps are performed and then we are going to build the machine learning model and will be deployed it on cloud platform.

2.4 Technical Requirements

Following are the requirements of this project:

• Model should be deployed on cloud (Azure, AWS, GCP, Heroku,Netlify).

• Cassandra database should be integrated in this project for any kind of user input.

2.5 Data Requirements

Data Requirement completely depend on our problem.

High Level Design

• For training and testing the model, we are using Campus Placement dataset which is available Kaggle.

• From user we are taking following input

Feature Names:

- Gender
- Percentage
- Qualification

Etc.

2.6 Tools Used

2.7 Data Requirements

- PyCharm is used as IDE.
- For visualization of the plots, Matplotlib, Seaborn are used.
- Horeku is used for deployment of the model.
- Front end development is done using HTML, CSS, Bootstrap, Flask is used
- for backend development and for API development.

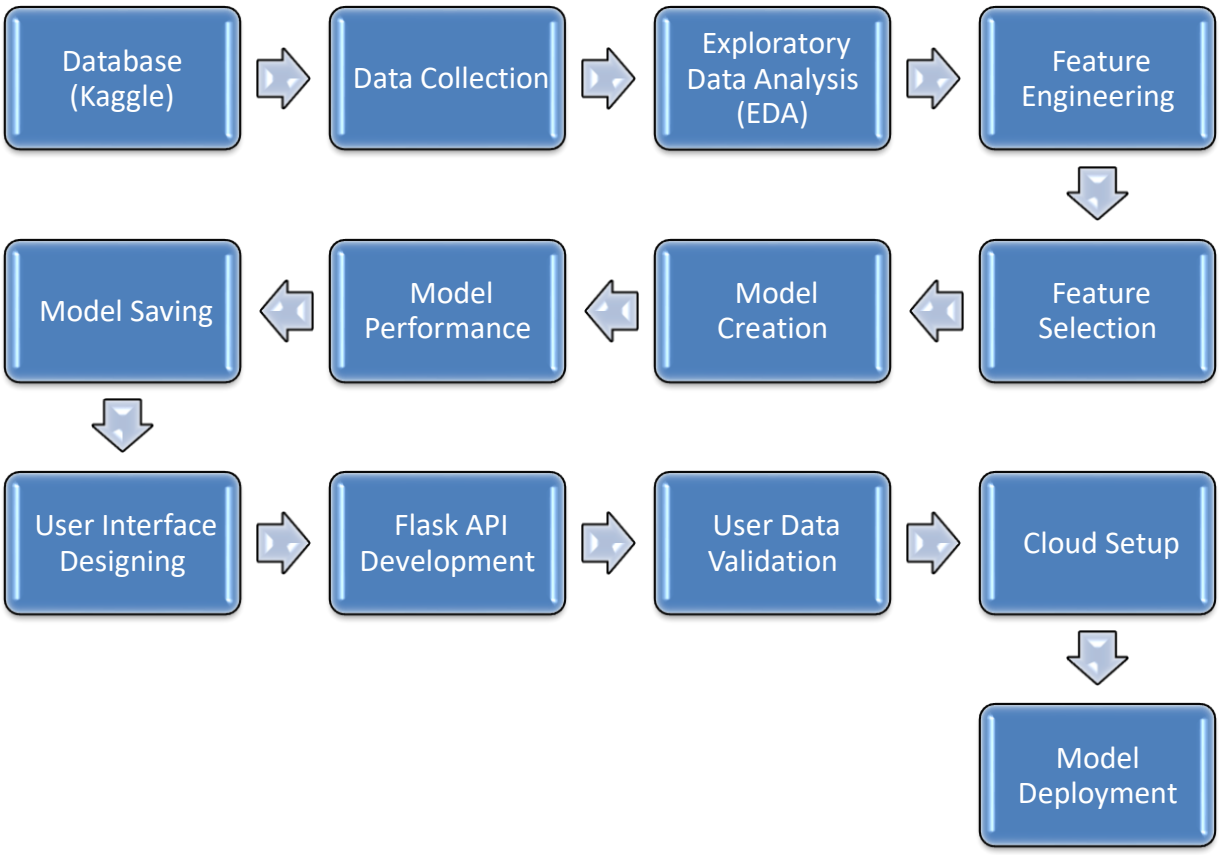
- GitHub is used as version control system

2.8 Constraints

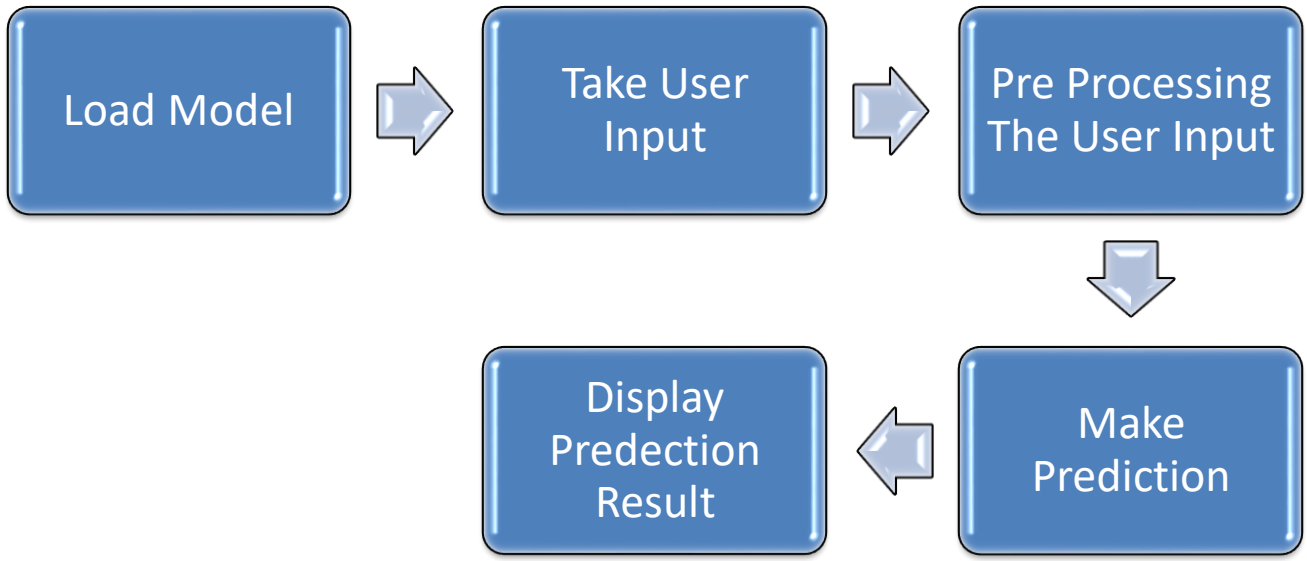
The Campus Placement Prediction Model system must be user friendly, errors free and users should not be required to know any of the back-end working

3. Design Details

3.1 Process Flow



3.2 Deployment Process



4. Performance

A. Solution of Campus Placement Prediction is used to predict the Student Placement , and it should be as accurate as possible.

B. That's why before building this model we followed complete process of Machine Learning. Here is summary of complete process:

i. First, we cleaned our dataset properly by removing all null value and duplicate value present in dataset.

ii. After that we performed EDA and feature transformation.

iii. And then we performed feature selection process.

iv. Then we performed the encoding – numerical features and categorical features

v. And now, we split the dataset in train-test split.

vi. After performing above, we trained our dataset on different classification algorithm (Logistic, SVM, KNN, Decision Tree Classifier, Random Forest Classifier etc.). After training the dataset on different algorithms, we got highest accuracy of 89.74% on SVM.

vii. After that we saved our model in pickle file format.

viii. After that our model was ready to deploy, we deployed this model on Horeku.

C. Re-usability

High Level Design

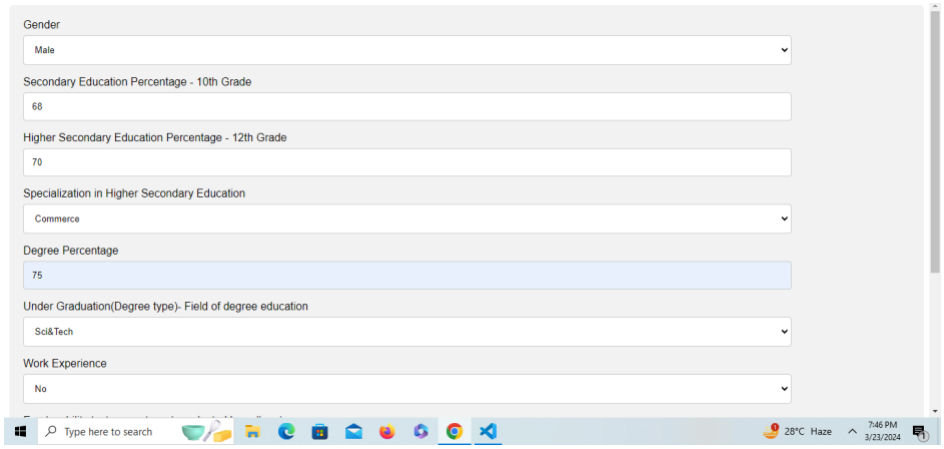
We have done programming of this project in such a way that it should be reusable. So that anyone can add and contribute without facing any problems

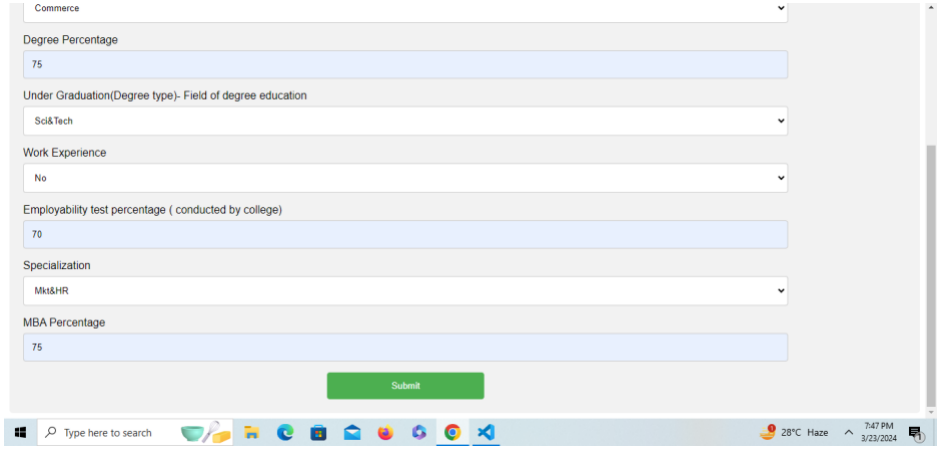
D. Application Compatibility

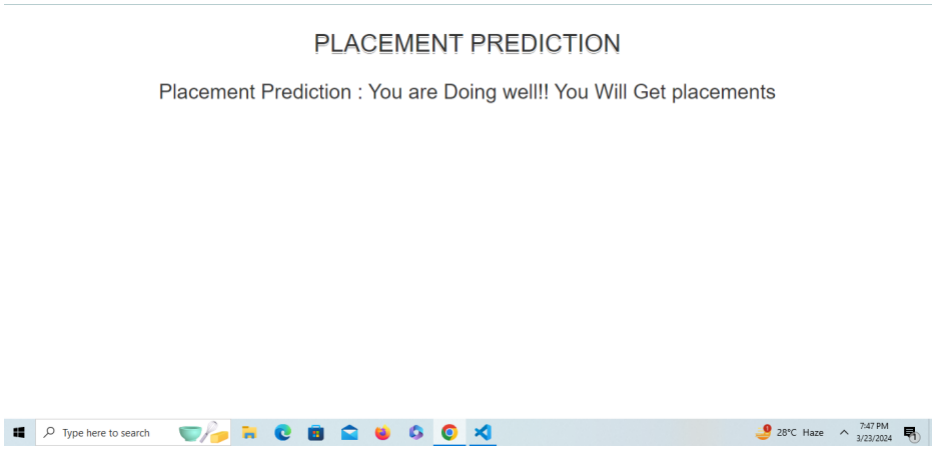
The different module of this project is using Python as an interface between them. Each module have it's own job to perform and it is the job of the Python to ensure the proper transfer of information.

E. Deployment: We have deployed this on Horeku cloud.

User Interface







5.Conclusion

The development and implementation of the machine learning campus placement prediction project hold significant promise for revolutionizing the recruitment process in educational institutions. Through the utilization of advanced algorithms and extensive data analysis, this project has demonstrated its capability to predict placement outcomes with a high degree of accuracy, thereby providing invaluable insights for both students and recruiters.

By leveraging historical placement data, academic performance metrics, and various other relevant factors, this project has not only facilitated informed decision-making for students regarding their career paths but has also empowered recruiters to streamline their selection processes and make more efficient hiring decisions.

Furthermore, the successful execution of this project underscores the immense potential of machine learning in addressing complex real-world challenges within the education and recruitment domains. As technology continues to evolve, there is boundless opportunity for further refinement and expansion of these predictive models, ultimately leading to enhanced efficiency, fairness, and transparency in the recruitment process.

In essence, the machine learning campus placement prediction project stands as a testament to the transformative impact of data-driven approaches in shaping the future of education and employment, offering a glimpse into a more streamlined and equitable recruitment landscape for students and employers alike.