



CAMPUS PLACEMENT PREDICTION

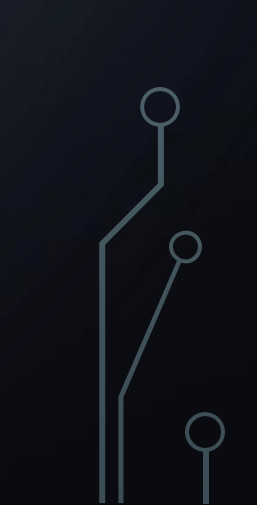

TECHNOLOGY-MACHINE LEARNING TECHNOLOGY

DOMAIN-HUMAN RESOURCE



PROBLEM STATEMENT:

The Placement of students is one of the most important objective of an educational institution. Reputation and yearly admissions of an institution invariably depend on the placements it provides it students with. That is why all the institutions, arduously, strive to strengthen their placement department so as to improve their institution on a whole. Any assistance in this particular area will have a positive impact on an institution's ability to place its students. This will always be helpful to both the students, as well as the institution.



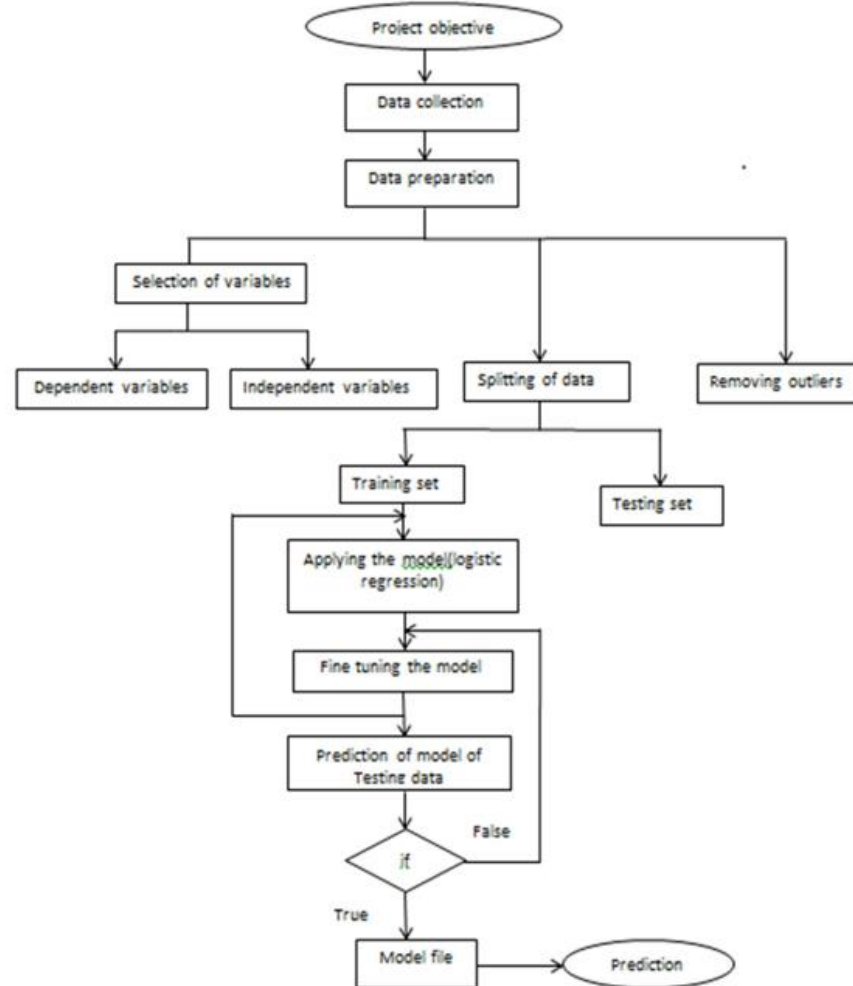
OBJECTIVE:

The main objective is to predict whether the student will be recruited in campus placements or not based on the available factors in the dataset.

APPROACH:

The classical machine learning tasks like Data Exploration, Data Cleaning ,Feature Engineering, Model Building and Model Testing. Try out different machine learning algorithms that's best fit for the above case.

ARCHITECTURE:



PROPOSED SYSTEM:

Firstly the project object is designed and the data is collected by the college, after collecting the data we prepare the data into training and testing datasets. The data so formed is separated into dependent variables and independent variables then the outliers are removed.

- 1) The training data is provided to the model design for the training results.
- 2) The model is fine tuned and the training data is further classified according to the parameters specified to the model.
- 3) After the training data is ready the testing dataset which was stored earlier is further processed to the model for the prediction.
- 4) If the testing values comes out to be true the values are finally set as the prediction.
- 5) If the testing values comes out to be false than the values are again sent back for the fine tuning of the model.

PROCEDURAL MODEL TRAINING:

❑ Data Export from Db :

The accumulated data from db is exported in csv format for model training.

❑ Data Preprocessing

Performing EDA to get insight of data like identifying distribution , outliers ,trend among data etc.

Check for null values in the columns. If present impute the null values.

Encode the categorical values with numeric values.

Perform Standard Scalar to scale down the values.

Clustering –

KMeans algorithm is used to create clusters in the preprocessed data. The optimum number of clusters is selected by plotting the elbow plot, and for the dynamic selection of the number of clusters, we are using KneeLocator function. The idea behind clustering is to implement different algorithms on the structured data.

The Kmeans model is trained over preprocessed data, and the model is saved for further use in prediction

Model Selection –

After the clusters are created, we find the best model for each cluster. By using 2 algorithms “SVM” and “XGBoost”. For each cluster both the hyper tuned algorithms are used. We calculate the AUC scores for both models and select the model with the best score. Similarly, the model is selected for each cluster. All the models for every cluster are saved for use in prediction

RESULTS:

- The proposed system results in the probability of the students getting placed in the campus drives. The 'CAMPUS PLACEMENT PREDICTION' provides the help for both students and the institution .The institution can focus on the potential students by knowing the prediction of this model. The technique we used is the SUPPORT VECTOR MACHINE[SVM] which gives the accuracy of students getting placed . The project mainly gives the information about the students probability of getting placed in the campus drives which benefits both the students and the institution.