Campus Recruitment

(Academic and Employability Factors influencing placement)

Submitted By : Sumith T S & Sreelekshmi S

Data Overview

This data set consists of Placement data of students in XYZ campus. It includes secondary and higher secondary school percentage and specialization. It also includes degree specialization, type and Work experience and salary offers to the placed students.Repository located at the following URL: https://www.kaggle.com/benroshan/factors-affecting-campus-placement .The Dataset consists of the qualitative and the quantitative parameters of the students that are necessary for improving the prediction.The most appropriate target would be statusvariable.

Goals

1. Which factor influenced a candidate in getting placed?
2. Which degree specialization is much demanded by corporate?
3. What is the Prediction percentage using differnt ML algorithms?
4. Build a model to predict if a student gets placed.

Specifications

1. Sl_no(Serial Number)
2. Gender- Male='M',Female='F'
3. Secondary Education percentage- 10th Grade(ssc_p):numerical
4. Board of Education- Central/ Others(ssc_b)
5. Higher Secondary Education percentage- 12th Grade(hsc_p):numerical
6. Board of Education- Central/ Others(hsc_b)
7. Specialization in Higher Secondary Education(hsc_s)
8. Degree Percentage(degree_p):numerical
9. Under Graduation(Degree type)- Field of degree education(degree_t)
10. Work Experience (workex)
11. Employability test percentage (conducted by college)(etest_p):numerical

Methodology

- Pre processing
  - Attribute selection
  - Cleaning missing values
  - Training and Test data
  - Feature Scaling

- Processing
  - Processing is applying different algorithms to the data to find the best results

Algorithms used

1. Logistic Regression
2. Random Forest Classifier
3. Support Vector Machine
4. XGBoost
5. Naive Bayes

Results

The data set used for is further splitted into two sets consisting of two third as training set and one third as testing set. Among the five algorithms applied XG Boost with hyper parameter tuning shown the best results.Accuracy is given by using following formula:

Accuracy= (TP+TN/TP+FN+FP+TN)∗100

where TP, TN, FN, FP represents the number of true positives, true negative, false negative and false positive cases.

Using XG Boost with hyperparameter tuning method :Accuracy 91.54%.

we can see the prediction and say that XG Boost with hyperparameter tuning model is better perform than other model.