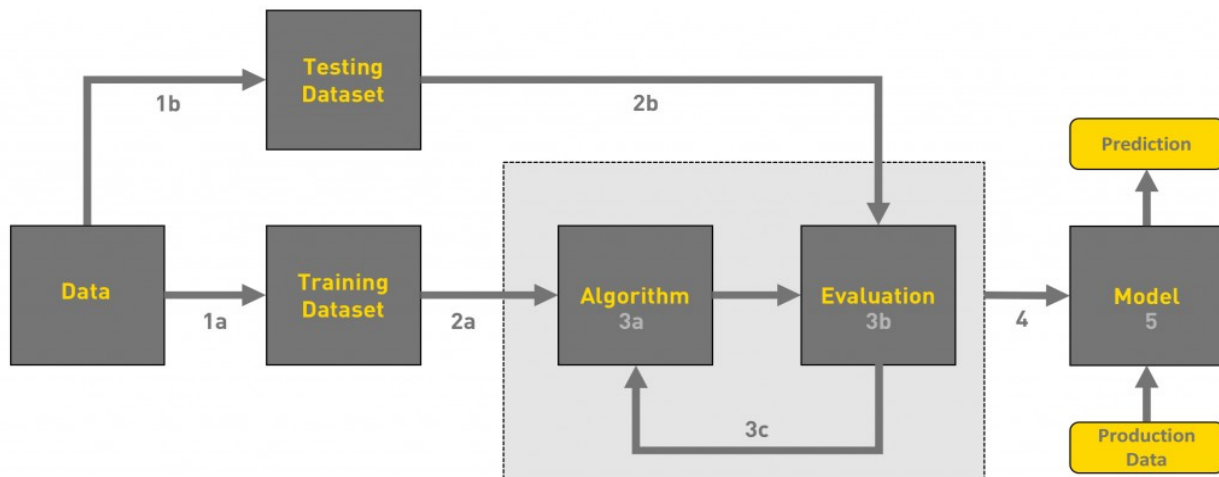


Predictive Attrition Model

Introduction

In this model, generally data has been prepared using googleform for being interviewed for Data Science and Data Analytics job considering different skills for this job and including experience and age of the candidates. This involves data cleaning, data visualisation and testing model using different machine algorithm as considering classification Problem and observing accuracy for every algorithm and predicted whether candidates are fair to be hired for company or not.



Breaking down the process of predictive model

- 1.Descriptive analysis of data-20% time
- 2.Data cleaning-20% time
- 3.Data Visualisation-40% time
- 4.Prediction-20% time

Data Analysis/Data Exploration

In this analysis the dataset involving different steps:-

- 1.showing the top head of the data, correlation of the data.
- 2.Described the data with maximum, minimum,mean, mode ,median and count.
- 3.Identify columns with maximum values.
- 4.Identify the outliers in dataset.

Data Cleaning

- 1.Filling the missing values with forward data.
- 2.Converting the bool values into integers with 0 and 1.
- 3.Filtered the data with different rating and created new data frame to further analysis.
- 4.Imputing the outliers values with mean and median.

Data Visualisation

- 1.plotting the bar graph with Experience and with their skill.
- 2.Plotting the subplots of graph comparing each skill with Other skill.
- 3.Plotting countplots for counting of number of candidate having skills and experience.

Machine Learning Algorithm

- 1.Splitting the dataset into two halves for Training set,Test set having 67% for Training set and 33% of Test Set.
- 2.Using different Classification algorithm I.e; Linear Regression.

Training and Testing the Model

We have calculate the accuracy of the every model Training and Testing the model For training and testing the data we have splitted data into two sets of ratio 67:33 and trained the data with 67% and tested with 33%.

After training we have tested the model and predicted the model with accuracy with approx 97%. This model predicts resume weightage of each candidate who applied for this job and we can also implement Clustering of words for achievements coloumns where candidates can fill about their internship,participations in hackathons or Google Summer of code on which we can implement machine Learning technique to fetch data.But here we have only decide the resume weightage on the basic of their skills,experience and CGPA.

Conclusion:-

In this model we have calculated whether the candidates resume weightage on the basic of their skills,experience and CGPA. We have passed the model through LinearRegression algorithm so that the employee can hired using supervised learning and in that using as classification problem and dealing with different algorithm.