

Summary

This project is for determining the potential customer who has more probability to get converted to buy the course and to make better decision to make the conversion rate better.

1. Preprocessing and Data Cleaning

The data set we got is full of null values as the form filled by the customer have not filled all the columns which are not mandatory so we treated them by removing which have more than 45% of null value and fill the null values by mode value of that column which have less than 45% missing value.

2. Exploratory Data Analysis

We done some analysis of the data. We perform univariate analysis and bivariate analysis to find the relation between the columns and how they distributed according to each other.

We also created dummy variable for the categorical column as we can build our model on numerical data only.

3. Splitting Dataset

We split the data set into 75% train set and 25% test set.

4. Model Building

We used logistic regression algorithm to model

5. Evolution

The model has the 86.94% of accuracy.