

# Summary

1) This analysis is done for X Education and to find ways to get more industry professionals to join their courses. The basic data provided gave us a lot of information about how the potential customers visit the site, the time they spend there, how they reached the site and the conversion rate.

2) In order to solve the problem statement and meet the specific business requirements we used the logistic regression model.

3) The following steps were used for model preparation

a) Data Cleaning and Exploratory Data Analysis

The data was partially clean except for a few null values and the option select had to be replaced with a null value since it did not give us much information. EDA was done to get the meaningful insights. There were a lot of elements in the categorical variables which were irrelevant. The numeric values seem good and less outliers were found.

b) Data Preparation:

Categorical variables were converted into numerical data using dummification of variables. The data was trained and splitted into the ratio of 70:30. Data points were standardised using MinMax scaler for numerical columns to avoid bias of some variables.

c) Model Building

RFE was done to attain the top 15 relevant variables. Later the rest of the variables were removed manually depending on the p-value (The variables with p-value < 0.05 were kept) on both train and test dataset.

d) Model evaluation

Confusion matrix preparation and using the optimum cut off value we checked the accuracy, sensitivity and specificity which came to be around 80% each. Prediction was done on the test data frame and with an optimum cut off as 0.46 with accuracy, sensitivity, specificity, precision and recall of 80~90%.

4) Conclusion and Recommendation

Features with good conversion rate are Tags\_Closed by Horizon, Tags\_Will revert after reading the mail, Total time spend on Website, Tags\_Others.

The people who are working professional, came to website by lead add form and have notable last activity could be the potential lead to be converted if proper counselling and information is provided about the platform.