Summary

The analysis is conducted for X Educational company to increase the conversion of the leads. The following steps were done:

1.Data cleaning: Removing unnecessary columns from the dataset. Select values were replaced as null values. Select values represent optional questions on the website. Dropping columns which has more than 35 percent null values.

2. EDA: Numeric values have no outlier present in the data.

3.Dummy variables: Dummy variables are created for categorical columns. Scaling we used MinMaxScaler.

4. Train test split: Splitting the data into train and test respectively.

5. Model building: Using RFE, we attain the top 15 variables. We remove the features based on high VIF and p-values.

6.Model Evaluation: Evaluation of model using accuracy, sensitivity and specificity came to around 80 percent.

7. Prediction: Prediction was done on the test data frame with optimal cutoff as 0.38.

8.precision-recall: precision was around 77 percent, and recall came to around 70 percent.

9.We can conclude by stating that the variables that highly matter in converting are:

    1.) 'The Total Time Spent on the Website'

    2.) 'Total Number of Visits'

    3.) 'When the lead source was Google'

    4.) 'When the last activity was SMS'

    5.) 'When the lead origin is Lead add format.'

    6.) 'When their current occupation is as a working professional.'

Thus, we can recommend the X Education to keep in mind the above-mentioned variables to enhance their success rate and increase the conversions as well.