

Lead scoring case study – Summary report

Based on the data provided, an analysis has been done on an online educational company to identify their potential leads and make sure right leads are converted to generate business and profits.

Following are the procedures that were followed to make sure detailed analysis is done and model is evaluated to reach a conclusion.

Data cleaning:

Data was cleaned efficiently. 'Select' category were replaced with Nan and ID columns were dropped. Further, variables with null values greater than 40% and variables with data variance were dropped.

Exploratory data analysis:

EDA was performed to visualize the data patterns and identify the behavior of variables in the given dataset.

Dummy Variables:

We created dummy variables and later scaling was done on continuous variables.

Train-Test split:

For train data and test data the split was done at 70% and 30% respectively.

Missing Value Imputation:

Null values have been replaced with a new data to make sure that we are not missing out on any valid data for analysis and further outlier treatment was done on continuous variables.

Dummy Variables:

We created dummy variables and later scaling was done on continuous variables.

Model Building:

Top 20 variables were identified by doing RFE. Later each of the variable was removed based on p-value and VIF value. Finally, top 17 variables were identified.

Model Evaluation:

ROC curve was leveraged to identify optimal cut-off and further final values were predicted based on our final model. Also accuracy, precision, recall & F1 score were calculate.

It was found that the variables that mattered the most in the potential buyers are:

1. When Lead is tagged as,
 - a. Closed by horizon
 - b. Lost to EINS
 - c. Will revert after reading email
 - d. Busy
 - e. Graduation in progress
 - f. Ringing
 - g. Switched off
2. The total time spend on the Website.
3. When the lead origin is Lead add form.
4. When the Last Notable activity
 - a. SMS Sent
 - b. Other
 - c. Email Opened
5. When the Last activity is Email bounced or Sms Sent.
6. When their current occupation is working professional or unemployed.
7. Number of page views per visit.

Keeping these in mind the X Education can flourish as they have a very high chance to get almost all the potential buyers to change their mind and buy their courses.