LEAD SCORE CASE STUDY

Tanmaya Raj Behera

PROBLEM STATEMENT

X Education is an organization which provides online courses for industry professional. The company marks its courses on several popular websites like google.

X Education wants to select most promising leads that can be converted to paying customers.

Although the company generates a lot of leads only a few are converted into paying customers, wherein the company wants a higher lead conversion. Leads come through numerous modes like email, advertisements on websites, google searches etc.

The company has had 30% conversion rate through the whole process of turning leads into customers by approaching those leads which are to be found having interest in taking the course. The implementation process of lead generating attributes are not efficient in helping conversions

BUSINESS GOAL

The company requires a model to be built for selecting most promising leads.

Lead score to be given to each leads such that it indicates how promising the lead could be. The higher the lead score the more promising the lead to get converted, the lower it is the lesser the chances of conversion

The model to be built in lead conversion rate around 80% or more.

STRATEGY

- Import data

- Clean and prepare the acquired data for further analysis

- Exploratory data analysis for figuring out most helpful attributes for conversion

- Scaling features

- Prepare the data for model building

- Build a logistic regression model

- Assign a lead score for each leads

- Test the model on train set

- Evaluate model by different measures and metrics

- Test the model on test set

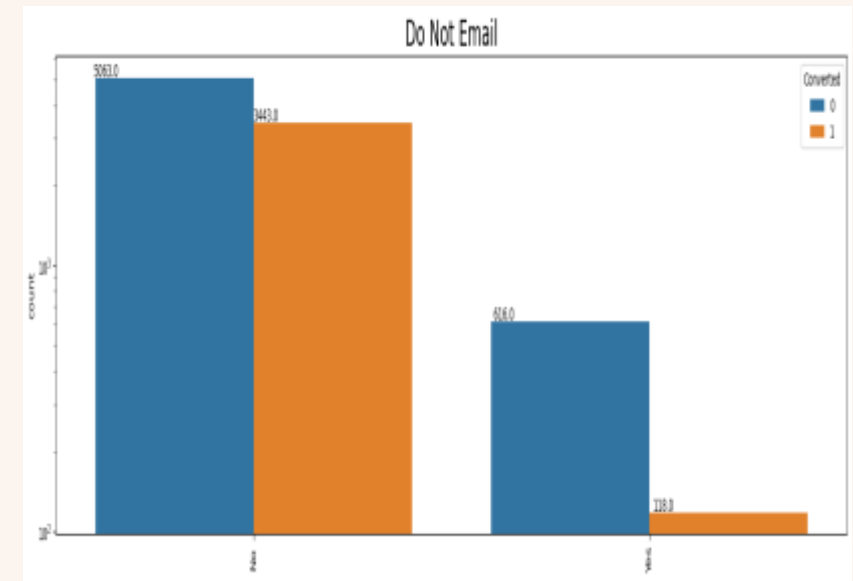- Measure the accuracy of the model and other metrics for evaluation

LEAD SOURCE VS CONVERTED

Google searches has had high conversions compared to other modes. Whilst references has had high conversion rate
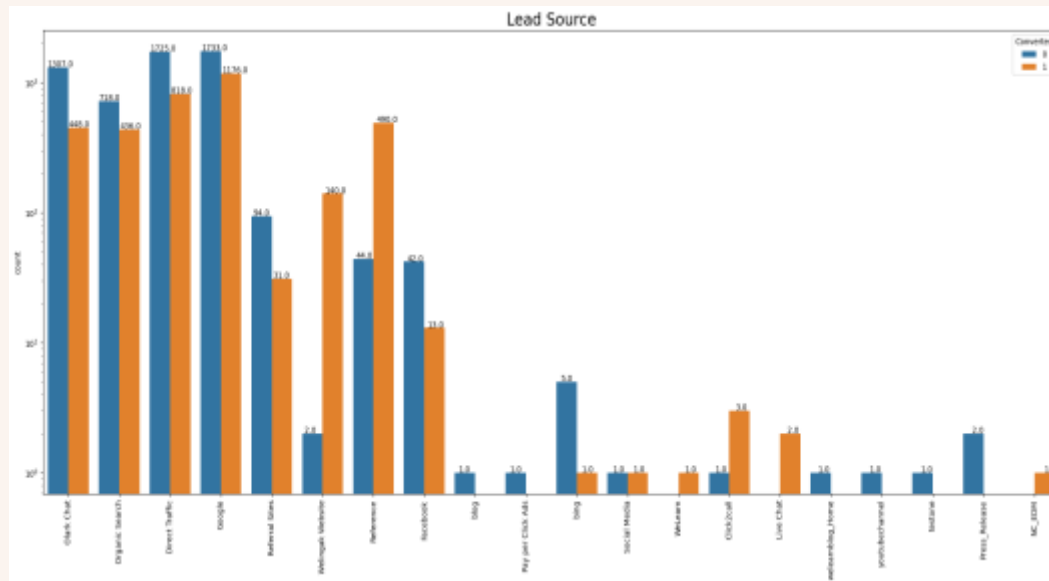
DO NOT EMAIL VS CONVERTED

Google searches has had high conversion compared to other modes. Whilst references has had high conversion rate
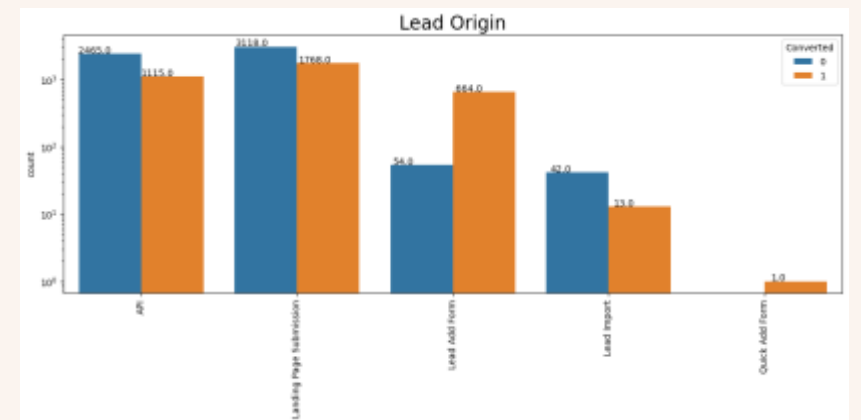
SPECILALIZATION VS CONVERTED

Most of leads have no information about specialization. Marketing management, human resources management has high conversion rate.
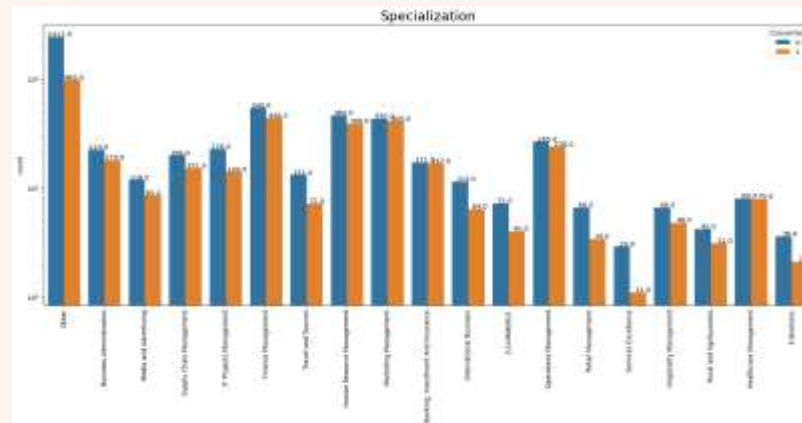
LEAD ORIGIN VS CONVERTED

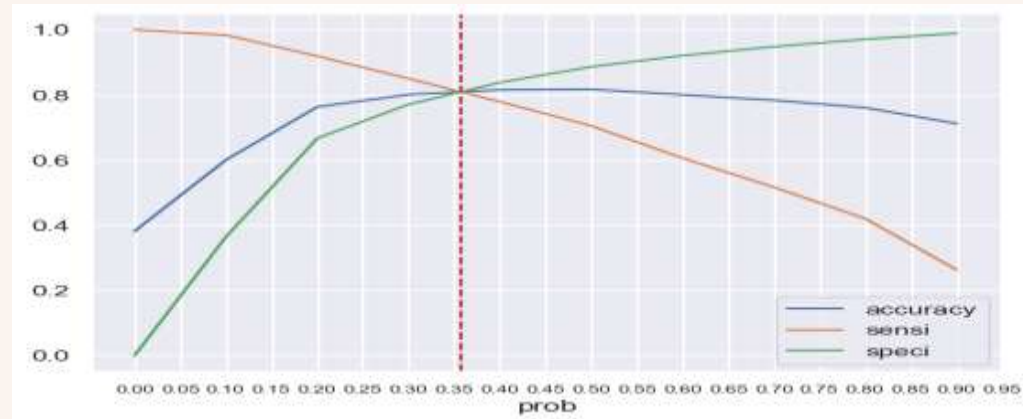Landing page submission has high lead conversions.

MODEL BUILDING

- Splitting into train and test set

- Scale variables in train set

- Build the first model

- Use RFE to eliminate less relevant variables

- Build the next model

- Eliminate variables based on high p-values

- Check VIF value for all the existing columns

- Predict using train set

- Evaluate accuracy and other metric

- Predict using test set

- Precision and recall analysis on test predictions

MODEL EVALUATION ON TRAIN

- 80.88% Accuracy
- 80.47% Sensitivity

- 72.47% Precision
- 80.47% Recall

CONCLUSI ON

EDA:

People spending higher than average time are promising leads, so targeting them and approaching them can be helpful in conversions

SMS messages can have a high impact on lead conversion

landing page submissions can help find out more leads

Marketing management, human resources management has high conversion rates. People from these specializations can be promising leads

References and offers for referring a lead can be good source for higher conversions An alert messages or information has seen to have high lead conversion rate

Logistic Regression Model:

The model shows high close to 80.88% accuracy

The threshold has been selected from Accuracy, Sensitivity, specificity measures and precision, recall curves.

The model shows 80.47% sensitivity.

The model finds correct promising leads and leads that have less chances of getting converted

Overall this model proves to be accurate