LEAD SCORING CASE STUDY

Presented By:

Tushar Parwani Tusarkanta Kar Tushar Geete

PROBLEM STATEMENT

Industry professionals can purchase online courses from X Education, an education firm.
 Many experts who are interested in the courses visit their website and look through the offerings on any given day.

Some of the leads convert during this process, but the majority do not. At X education,

the lead conversion rate is typically 30%.

You have been assigned by X Education to assist them in identifying the most promising prospects—that is,

the leads with the highest likelihood of becoming paying clients.

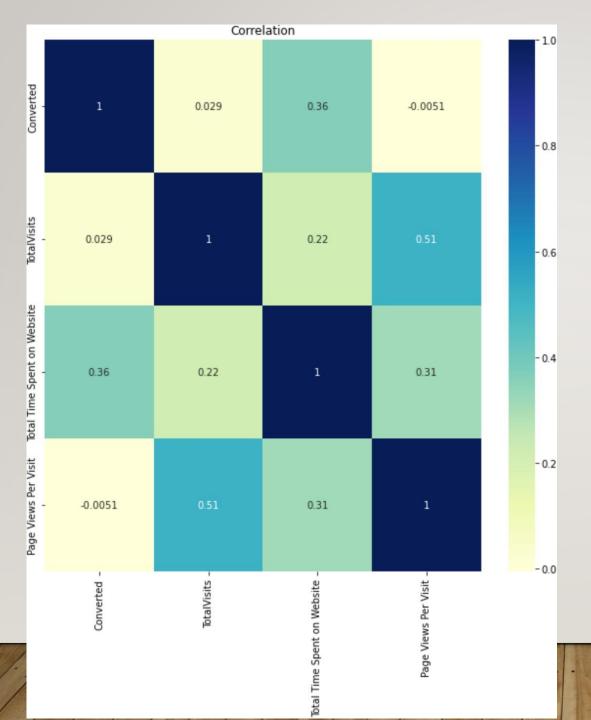
BUSINESS GOAL

 Your job at the company is to create a model in which you have to give each lead a score such that customers who have higher lead scores are more likely to convert, and customers who have lower lead scores are less likely to do so.

The intended lead conversion rate, as stated by the CEO in particular, is approximately 80%.

STRATEGY

- Import dataset
- Data Cleaning
- Exploratory Data Analysis
- Scaling Features
- Prepare the data for model building
- Logistic Regression Model
- Assign a lead score for each leads
- Train and test Model
- Evaluate the model
- Test the Model in test set
- Measure the accuracy of the model



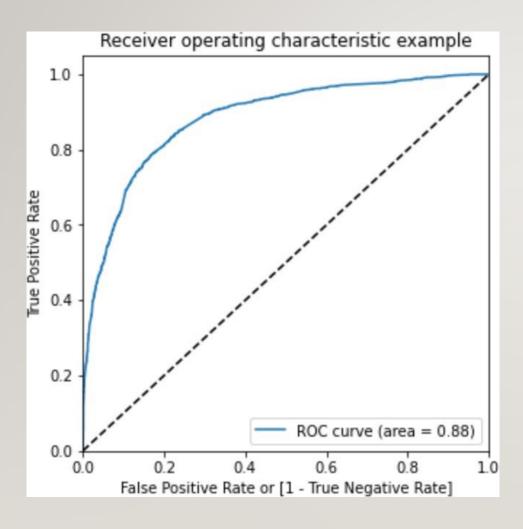
Correlation

- Total Visits
- Converted
- Total time spent on website

HEAT MAP TO CHECK FOR CORRELATION

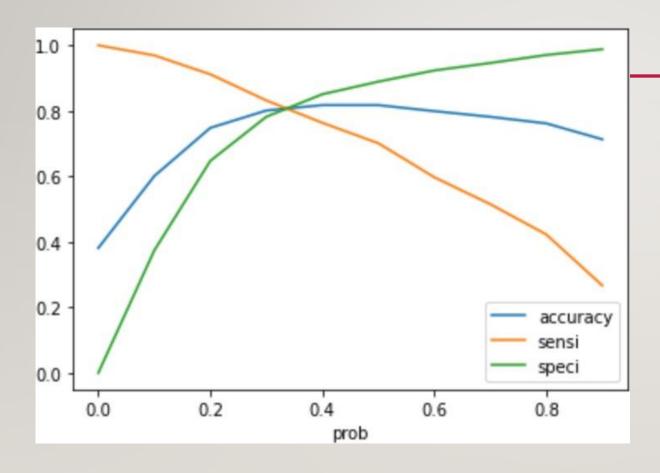
-1.00 Total Time Spent on Website X Education Forums - 0.75 Digital Advertisement A free copy of Mastering The Interview Lead Origin Lead Add Form Lead Origin Quick Add Form Lead Source_Google - 0.50 Lead Source Organic Search Lead Source Reference Lead Source Welingak Website Last Activity Email Link Clicked 0.25 Last Activity Form Submitted on Website Last Activity_Others Last Activity Last Activity_SMS Sent Last Activity Unsubscribed What is your current occupation Unemployed Last Notable Activity Email Link Clicked Last Notable Activity Modified Last Notable Activity Other Activity Last Notable Activity SMS Sent -0.25Specialization_Business Administration Specialization_E-COMMERCE Specialization Healthcare Management Specialization Human Resource Management Specialization International Business -0.50Specialization Media and Advertising Specialization Retail Management Specialization Services Excellence Specialization Travel and Tourism City Other Cities of Maharashtra City Thane & Outskirts

ROC Curve



Optimal cutoff probability is that probwhere we get balanced sensitivity and specificity.

MODEL EVALUATION(TRAIN)C



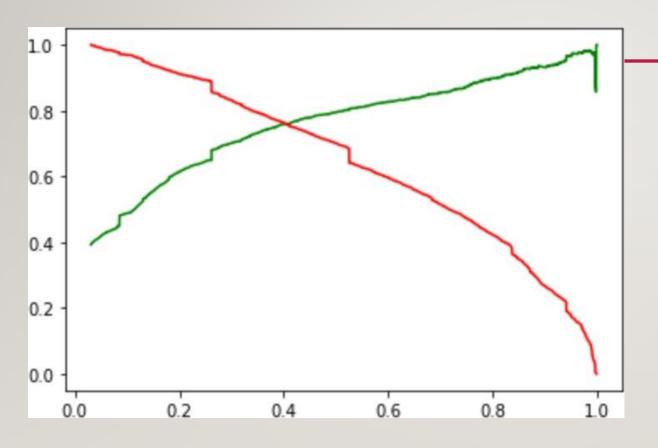
Accuracy sensitivity and specificity

Train Data:

Accuracy: 81.07% Sensitivity: 79.3%

Specificity: 82.1%

MODEL EVALUATION (TEST)



Accuracy sensitivity and specificity

Test Data:

Accuracy: 79.68%

Sensitivity: 81%

Specificity: 78.47%

CONCLUSION

The model seems to be performing well.

Train Data:

Accuracy: 81.07% Sensitivity: 79.3% Specificity: 82.1%

Test Data:

Accuracy: 79.68% Sensitivity: 81%

Specificity: 78.47%

The model appears to be operating efficiently. Is it possible to provide this model for making wise decisions.

Thank You.