Lead Score Case Study

SASIDHAR REDDY

Problem Statement

X Education sells online courses to industry professionals. The company markets its courses on several websites and search engines like Google.

Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos. When these people fill up a form providing their email address or phone number, they are classified to be a lead. Moreover, the company also gets leads through past referrals.

Once these leads are acquired, employees from the sales team start making calls, writing emails, etc. Through this process, some of the leads get converted while most do not. The typical lead conversion rate at X education is around 30%.

Business Goal:

X Education needs help in selecting the most promising leads, i.e., the leads that are most likely to convert into paying customers.

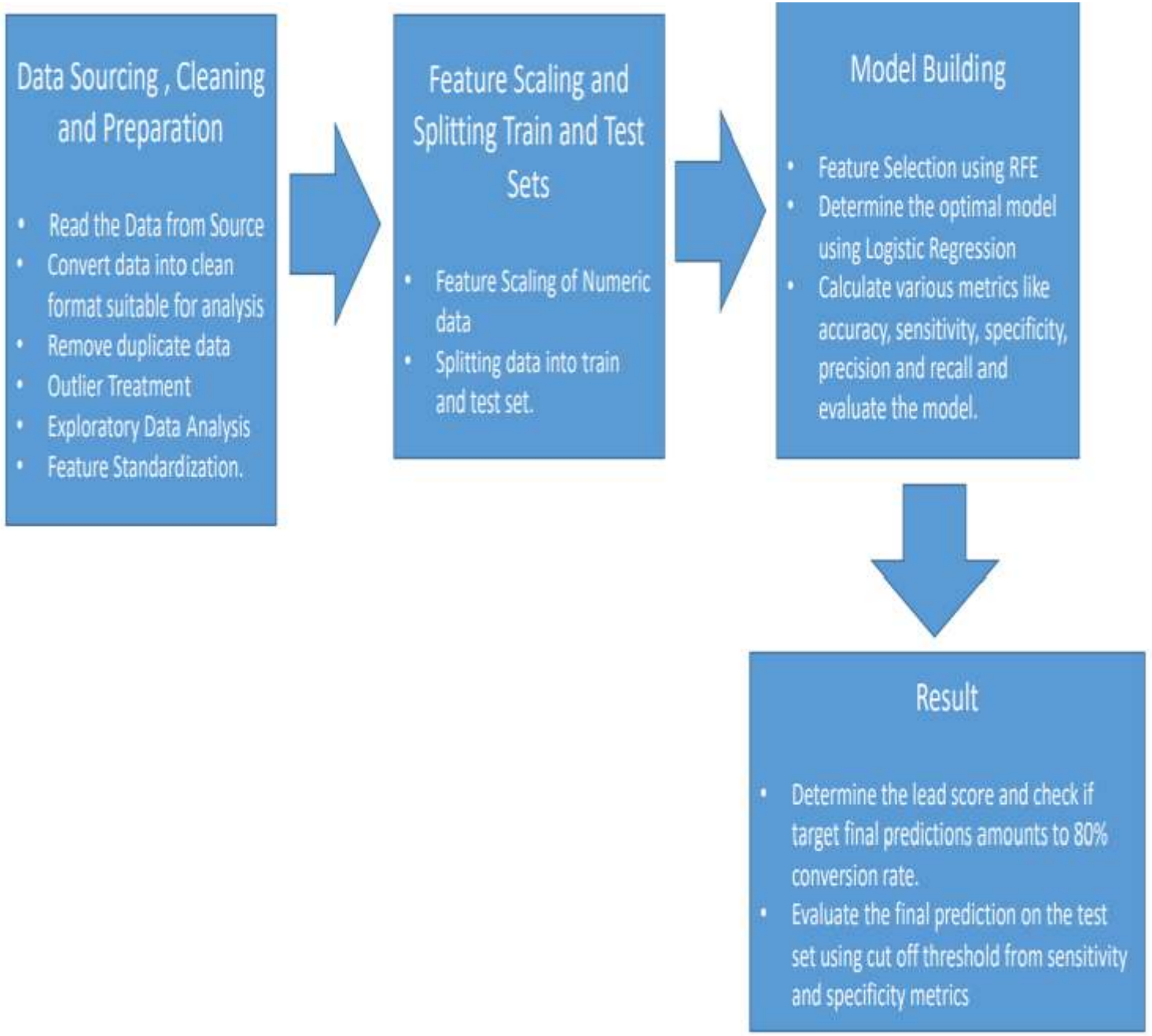
The company needs a model wherein you a lead score is assigned to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance.

The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%

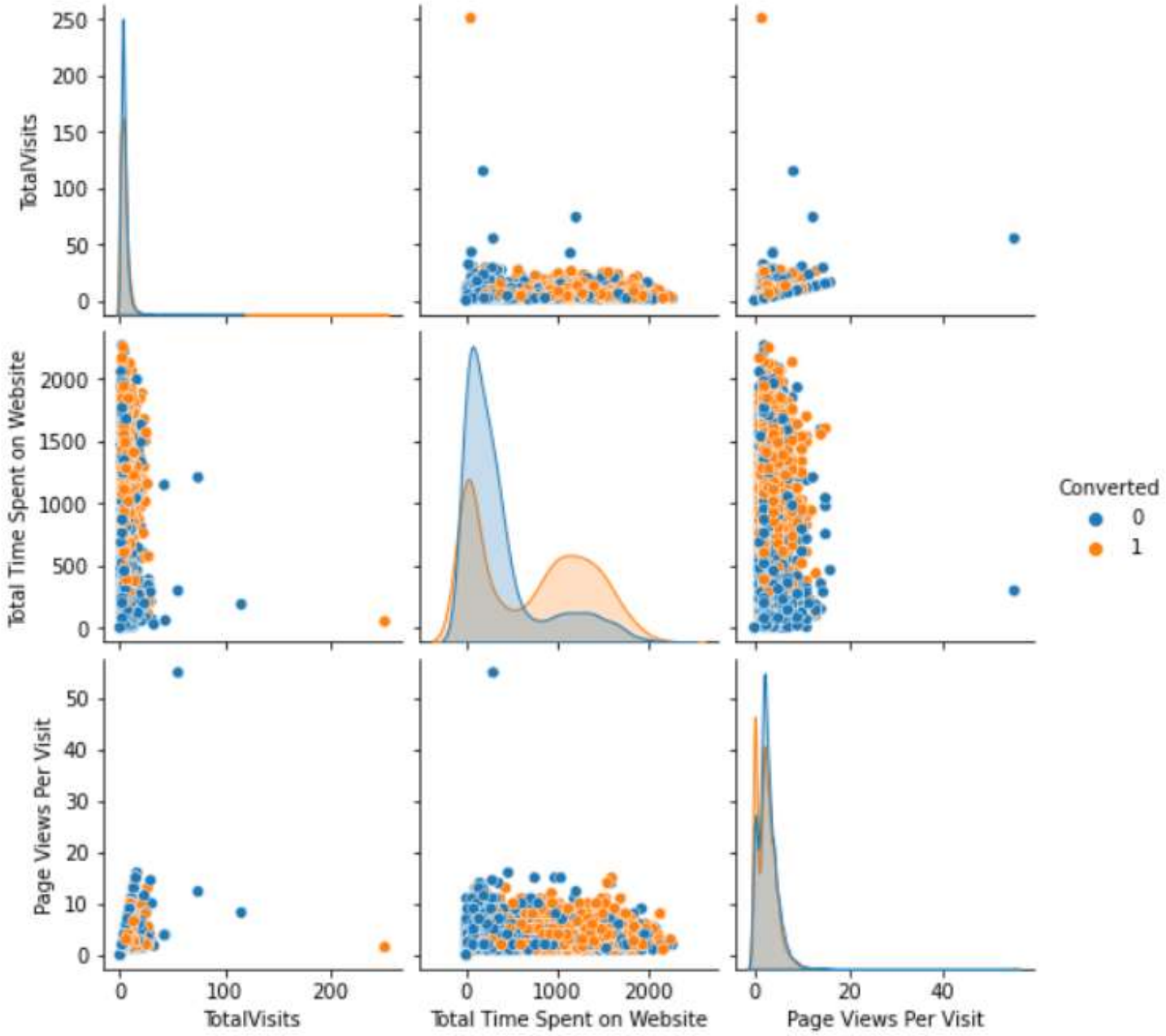
Strategy

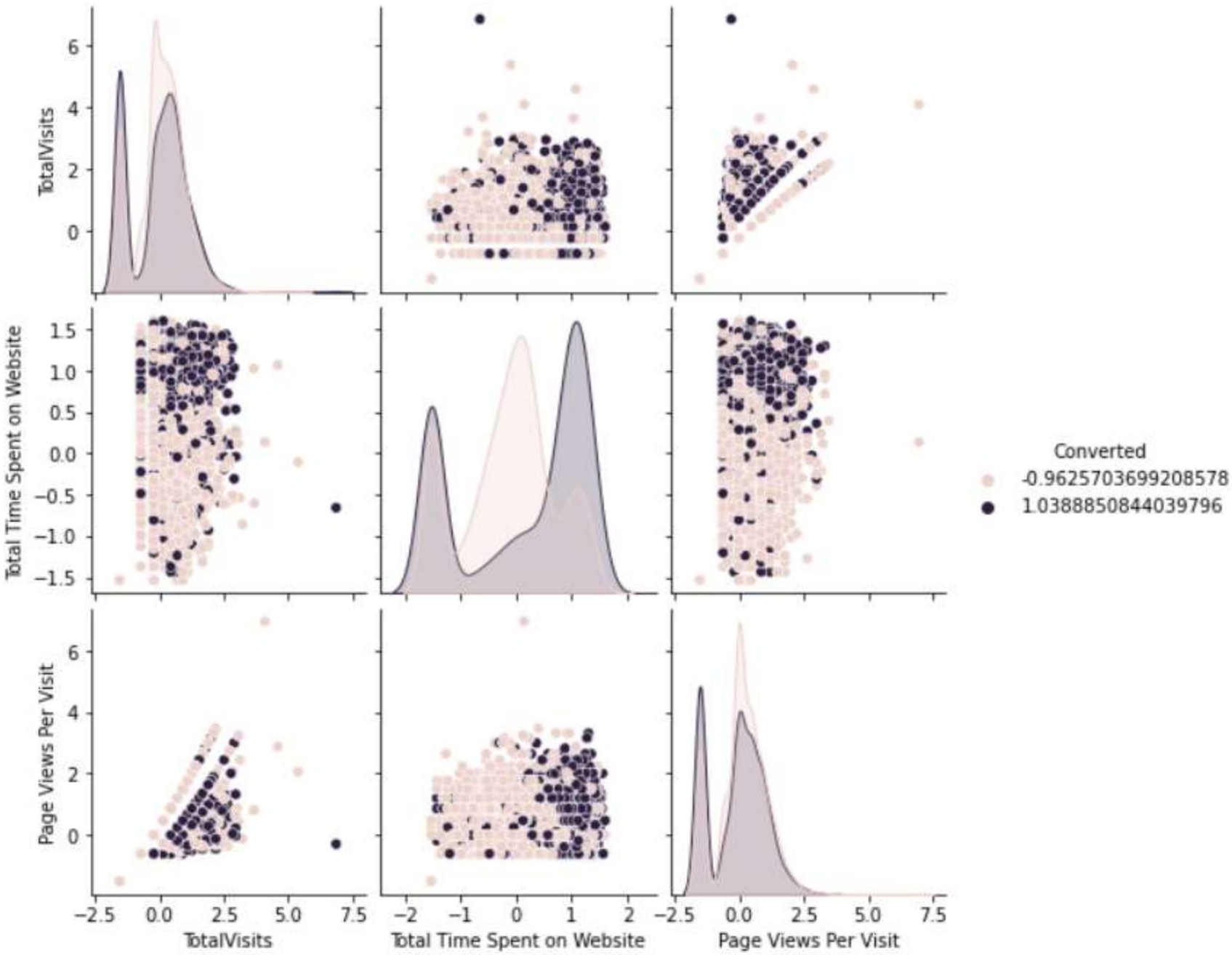
- Source the data for analysis
- Clean and prepare the data
- Exploratory Data Analysis.
- Feature Scaling
- Splitting the data into Test and Train dataset.
- Building a logistic Regression model and calculate Lead Score.
- Evaluating the model by using different metrics - Specificity and Sensitivity or Precision and Recall.
- Applying the best model in Test data based on the Sensitivity and Specificity Metrics.

Problem Solving Methodology



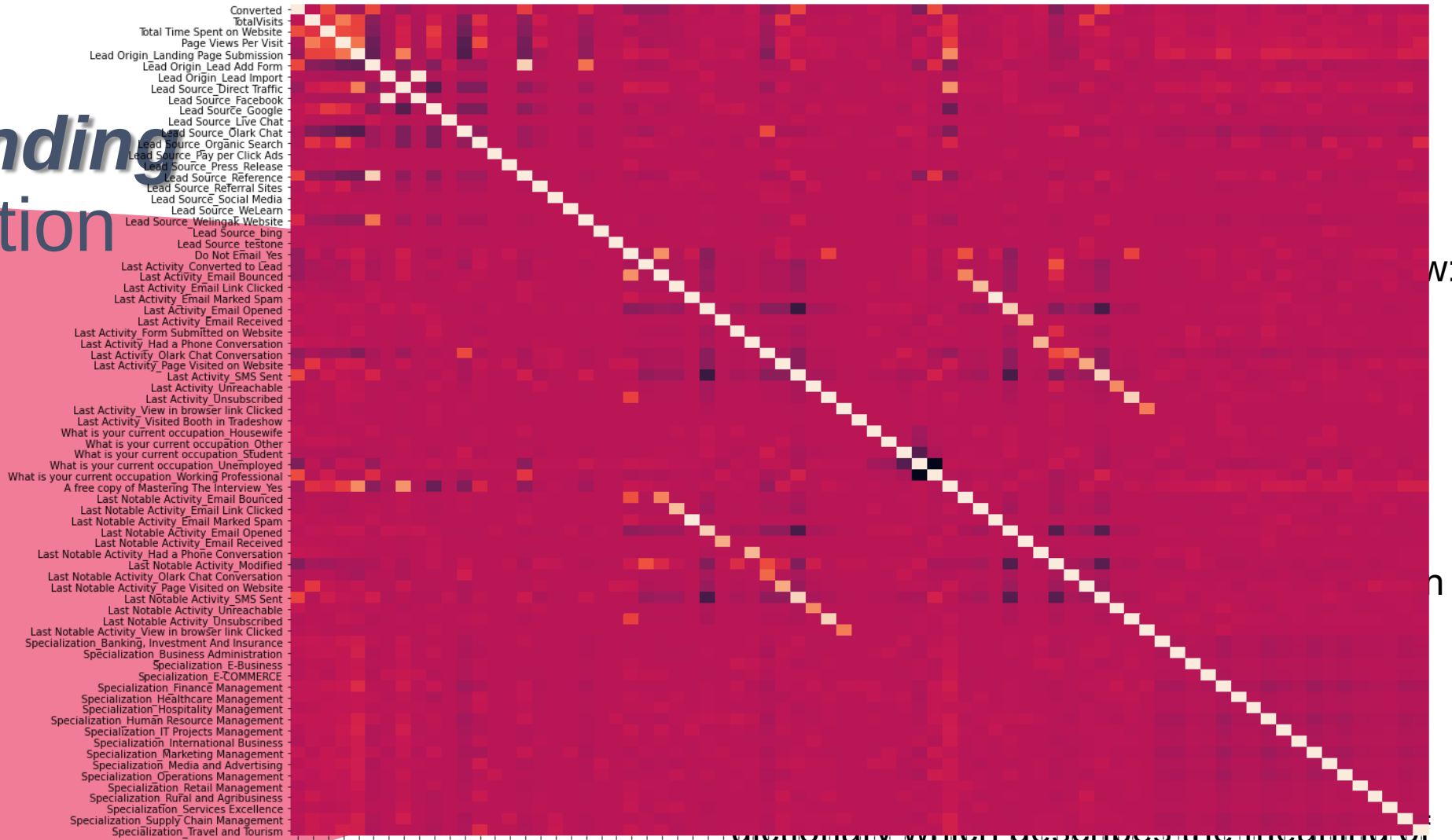
Exploratory Data Analys




TotalVisits
Total Time Spent on Website
Page Views Per Visit
TotalVisits
Total Time Spent on Website
Page Views Per Visit
Converted
-0.9625703699208578
1.0388850844039796

Data Understanding

by Correlation Table

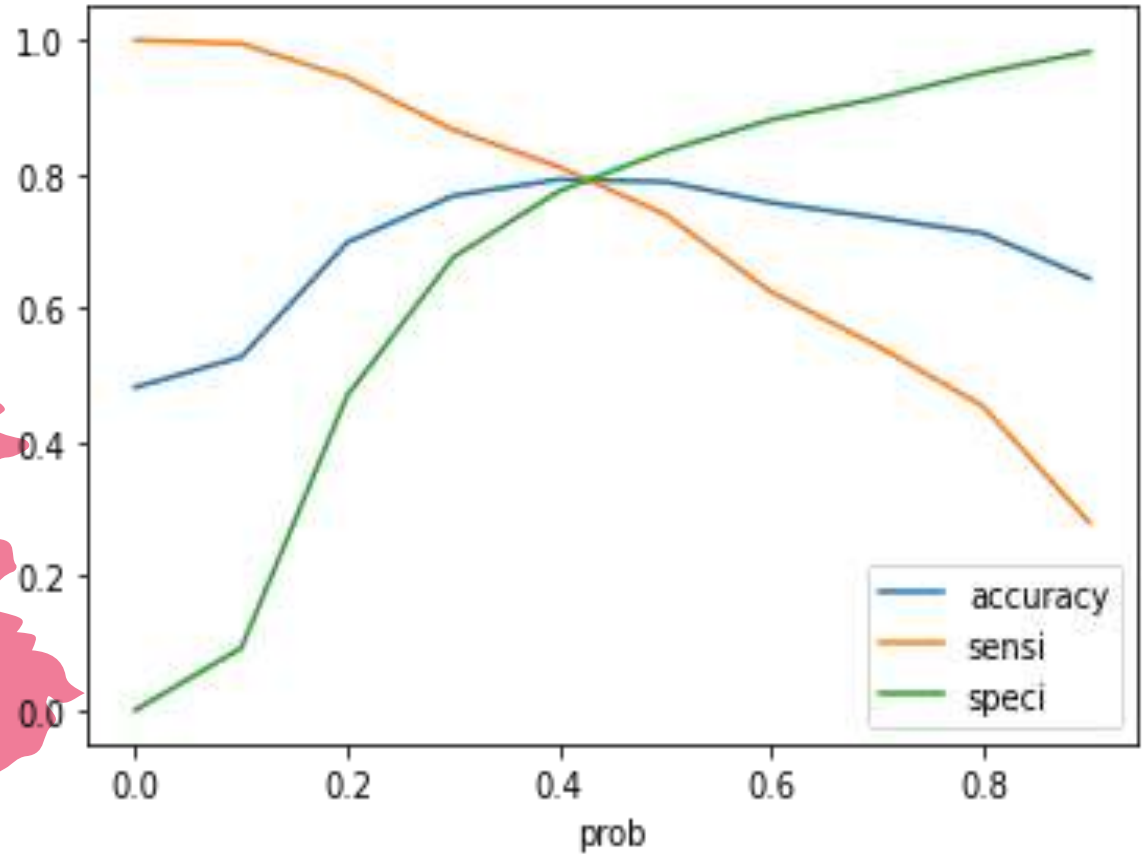


w:

n

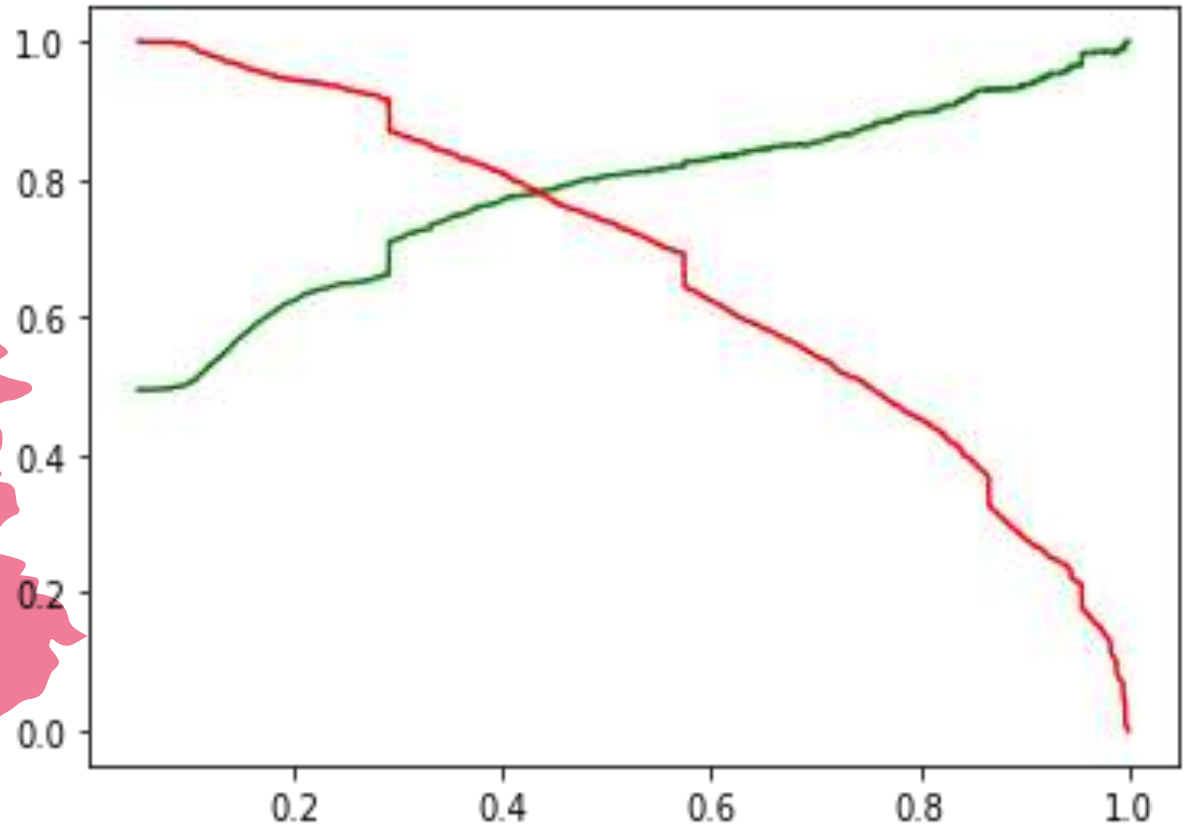
the variables

Model Evaluation - Sensitivity and Specificity on Train Data Set



As we can see that around 0.42, you get the optimal values of the three metrics. So let's choose 0.42 as our cutoff

Model Evaluation- Precision and Recall on Train Dataset



Precision - 78 % • Recall - 71 %

Model Evaluation – Sensitivity and Specificity on Test Dataset

Accuracy: 78%

Precession: 78%
Recall : 76%

Summary:

There are a lot of leads generated in the initial stage (top) but only a few of them come out as paying customers from the bottom. In the middle stage, you need to nurture the potential leads well (i.e., educating the leads about the product, constantly communicating etc.) in order to get a higher lead conversion. First, sort out the best prospects from the leads you have generated. 'TotalVisits' , 'Total Time Spent on Website' , 'Page Views Per Visit' which contribute most towards the probability of a lead getting converted. Then, You must keep a list of leads handy so that you can inform them about new courses, services, job offers and future higher studies. Monitor each lead carefully so that you can tailor the information you send to them. Carefully provide job offerings, information or courses that suits best according to the interest of the leads. A proper plan to chart the needs of each lead will go a long way to capture the leads as prospects. Focus on converted leads. Hold question-answer sessions with leads to extract the right information you need about them. Make further inquiries and appointments with the leads to determine their intention and mentality to join online courses.

Thank you

SASIDHAR REDDY K

vinsashi@gmail.com