Lead Scoring : By Mahesh, Nalini & Sushant

Problem Statement

AN EDUCATION COMPANY NAMED X EDUCATION SELLS ONLINE COURSES TO INDUSTRY PROFESSIONALS.

NOW, ALTHOUGH X EDUCATION GETS A LOT OF LEADS, ITS LEAD CONVERSION RATE IS VERY POOR OF ABOUT 30%
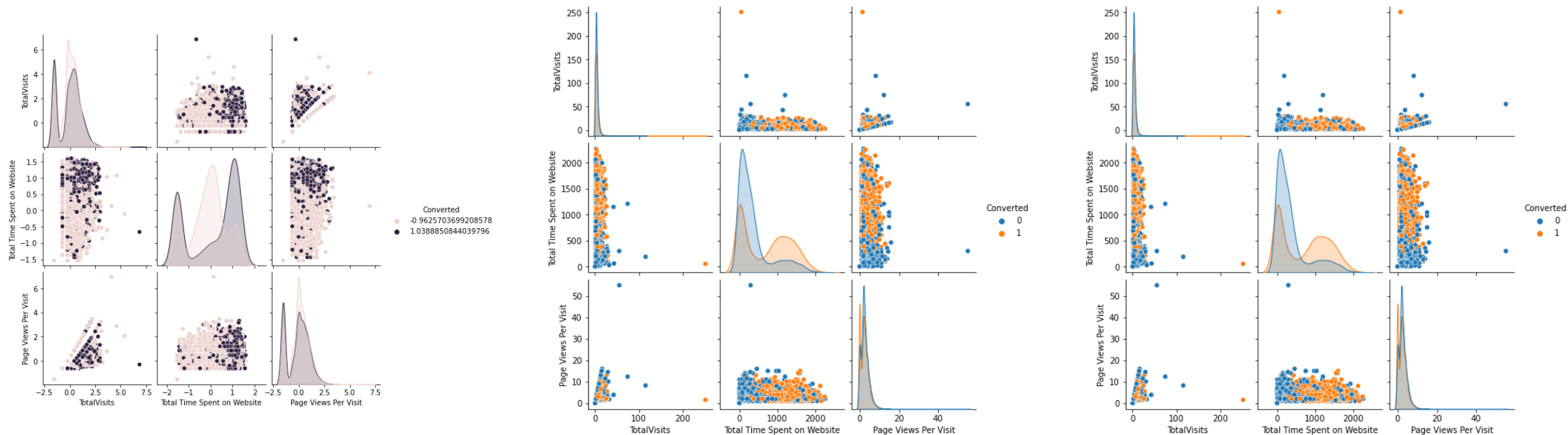
THE COMPANY WANTS TO INCREASE IT TO 80%

Build a logistic regression model to assign a lead score between 0 and 100 to each of the leads which can be used by the company to target potential leads.
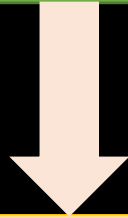
A higher score would mean that the lead is hot, i.e. is most likely to convert whereas a lower score would mean that the lead is cold and will mostly not get converted
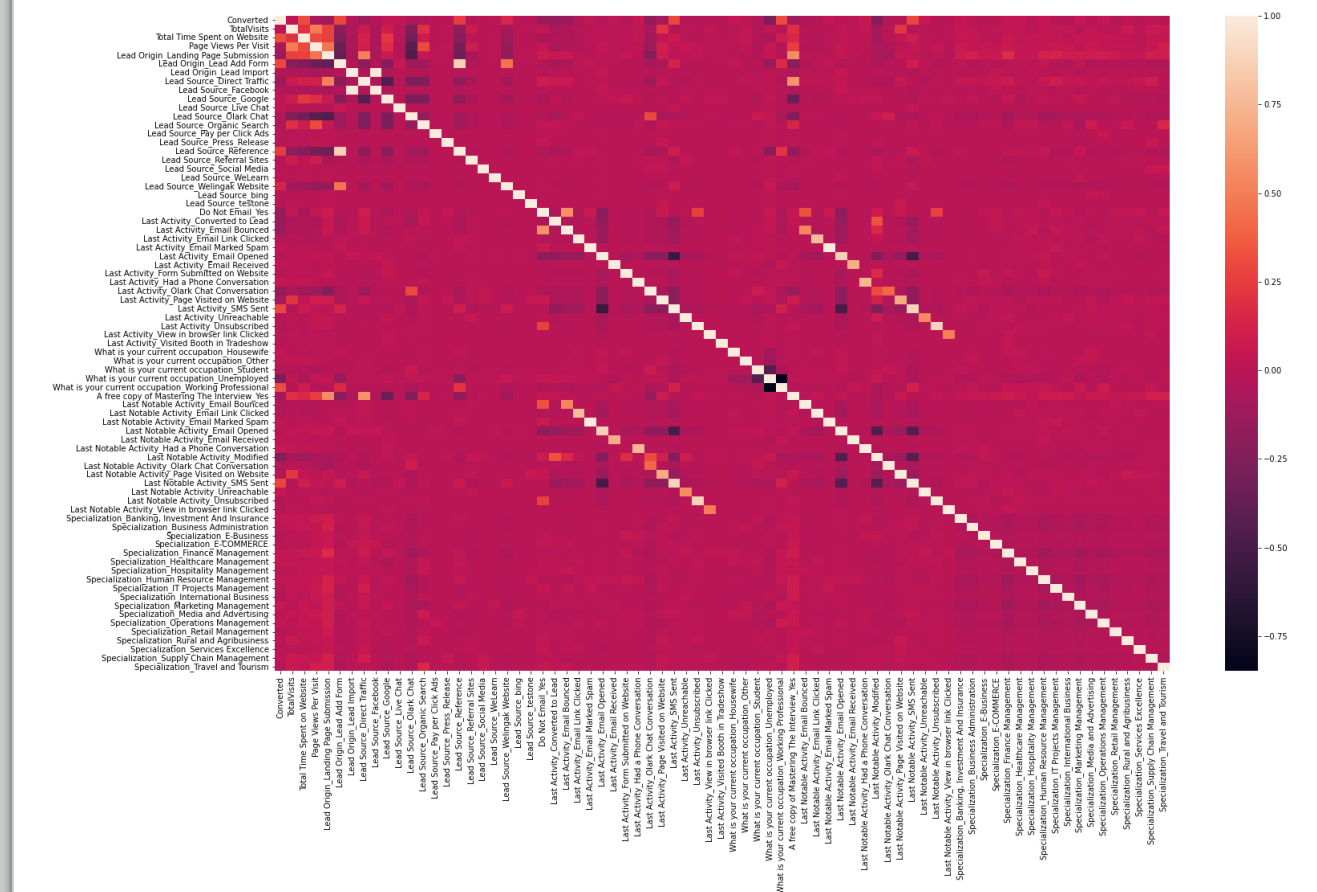
Goal

Preparing the Data for Modelling

Scaling :

We notice a few numeric variables in the dataset which have different scales. Hence, proceeding with scaling the variables
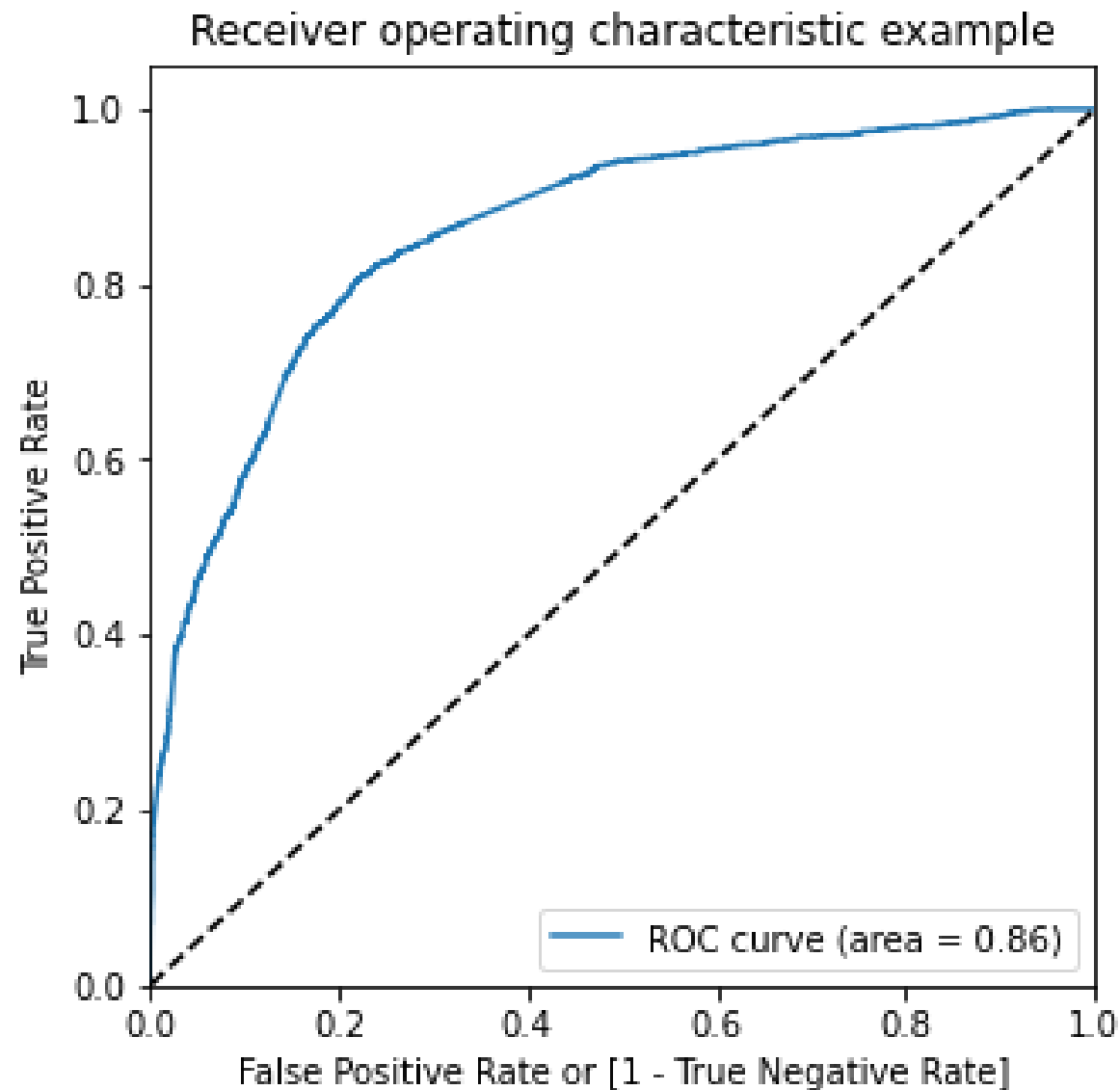
Scale the three numeric features present in the dataset.

# Looking at the correlation table ------->

Finding the Optimal Cut-off

The area under the curve of the ROC is 0.86 which is quite good. So we seem to have a good model. Let's also check the sensitivity and specificity tradeoff to find the optimal cutoff point.

Receiver operating characteristic example

To see the values of accuracy, sensitivity, and specificity at different values of probability cut-offs

As you can see that around 0.42, you get the optimal values of the three metrics. So let's choose 0.42 as our cut-off now.
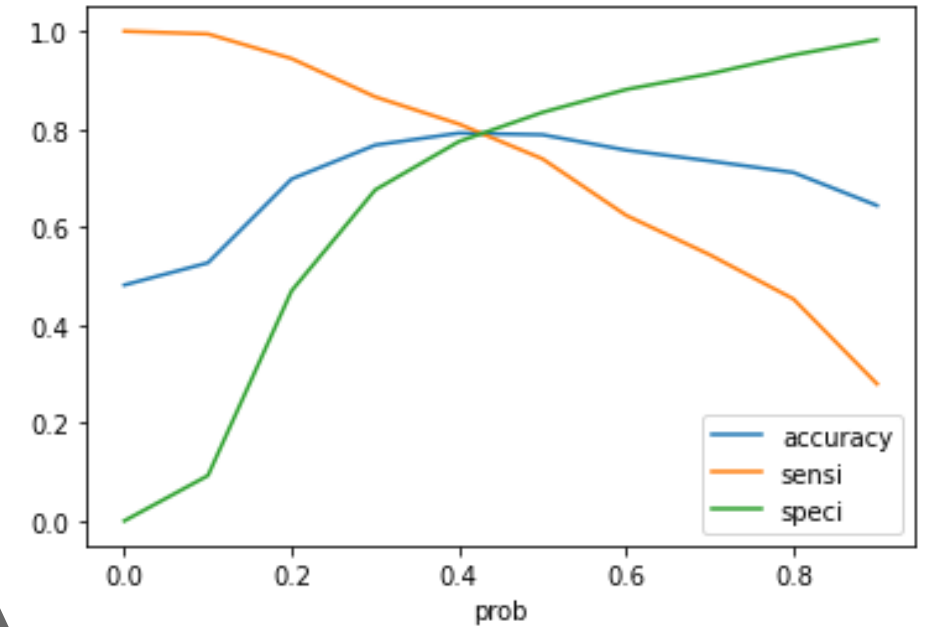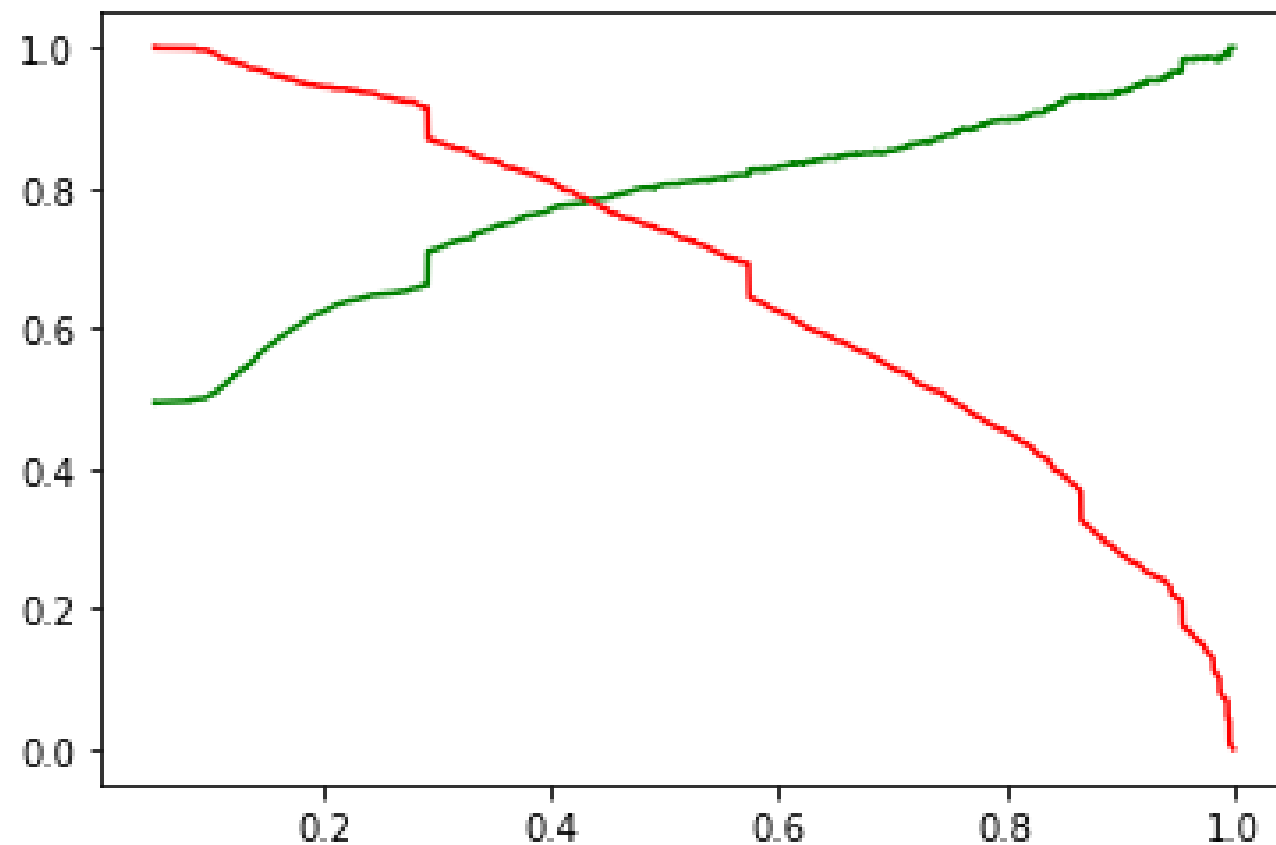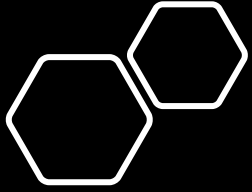
Precision and recall Trade-off

Calculate Recall

TP/(TP+FN) = `0.7771056305258259`

Cut-off points is good!!

Conclusion:

Lots of leads generated initially but only a few come out as paying customers by the time the analysis is completed

Potential leads needs to be nurtured in order to get higher lead conversion

'TotalVisits', 'Total Time Spent on Website', 'Page Views Per Visit' contribute most towards the probability of lead getting converted

List of leads must be kept handy to inform them about new courses, services, job offers and future higher studies

Monitor each lead and provide job offerings, information or courses that suits best according to the interest of the leads

Focus on converted leads

Further enquiries and appointments with the leads can be made to determine their intention to join the online courses