AN ORGANISATION CALLED X EDUCATION OFFERS ONLINE TRAINING FOR BUSINESS PROFESSIONALS. THE COMPANY LISTS ITS COURSES ON A NUMBER OF WELL-KNOWN WEBSITES, INCLUDING GOOGLE.THE MOST PROMISING LEADS THAT CAN BECOME PAYING CLIENTS SHOULD BE CHOSEN BY X EDUCATION.EVEN WHILE THE BUSINESS GENERATES A LARGE NUMBER OF LEADS, ONLY A SMALL PERCENTAGE OF THOSE LEADS END UP BECOMING PAYING CUSTOMERS. NUMEROUS CHANNELS, INCLUDING EMAIL, WEBSITE ADVERTS, GOOGLE SEARCHES, ETC., ARE USED TO GENERATE LEADS.BY ENGAGING LEADS WHO ARE KNOWN TO BE INTERESTED IN TAKING THE COURSE, THE ORGANISATION HAS EXPERIENCED A 30% CONVERSION RATE FOR THE ENTIRE PROCESS OF CONVERTING LEADS INTO CLIENTS. IMPLEMENTING LEAD-GENERATING QUALITIES IS INEFFECTIVE AT ASSISTING CONVERSIONS.

MD Abubakar

OBJECTIVES OF THE CASE STUDY

- This case study has a lot of objectives.
- Create a logistic regression model to provide each lead a lead score between 0 and 100 that the business may use to target potential prospects.
- In contrast, a lower number would indicate that the lead is chilly and unlikely to convert, while a higher score would indicate that the lead is hot and most likely to convert.
- You will also need to deal with some additional issues that the firm has raised and that your model should be prepared to address if the company's requirements change in the future.

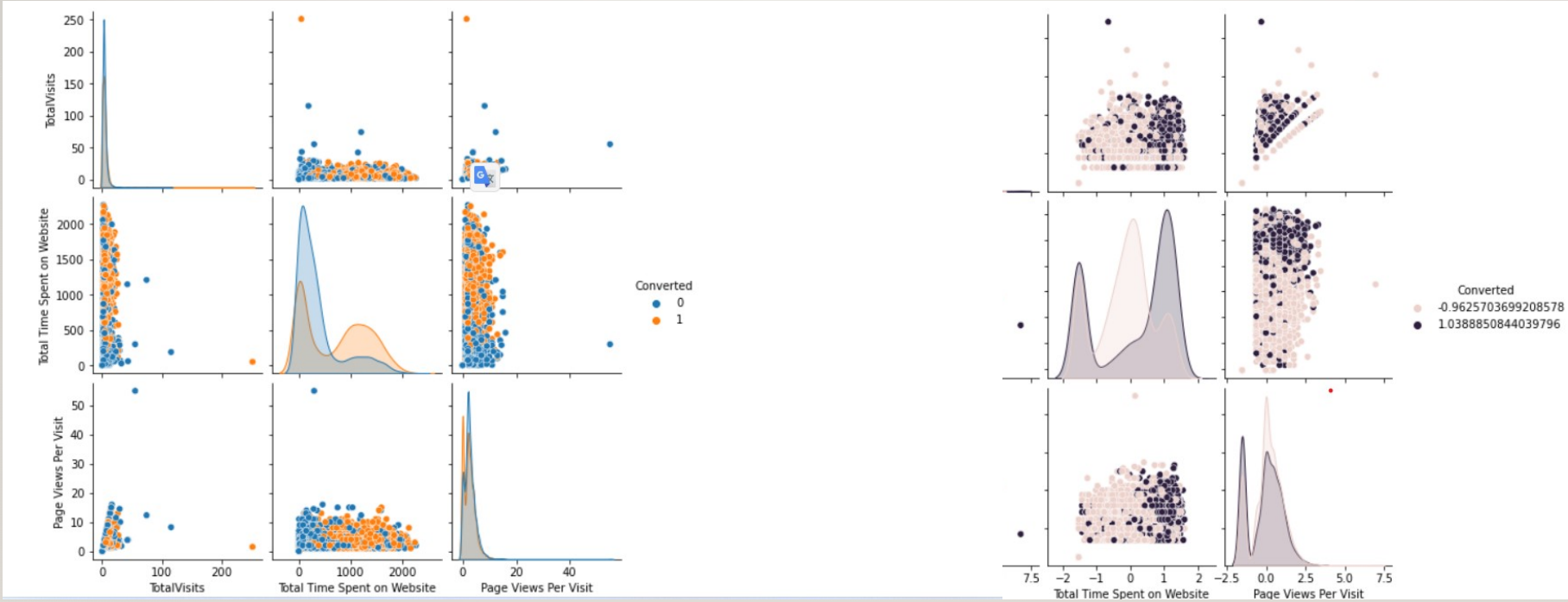
APPROACH

- Source the data For analysis
- Reading & Understanding the data
- Data Cleaning
- EDA
- Feature scaling
- Splitting the data into test & train dataset
- Prepare the data for modelling
- Model building
- Model evaluation-specificity & sensitivity or precision recall
- Making predictions on the test set

DATA SOURCING, CLEANING & PREPARATIONREAD

- the data from CSV File
- Outlier treatment
- Data cleaning -Handling Null Values & removing higher Null values data
- Removing Redundant columns in the data
- Imputing Null Values
- Exploratory data analysis-approx.
- Conversion Rate is 38%
- Feature standardization

GET THE DATA READY FOR MODELLING.

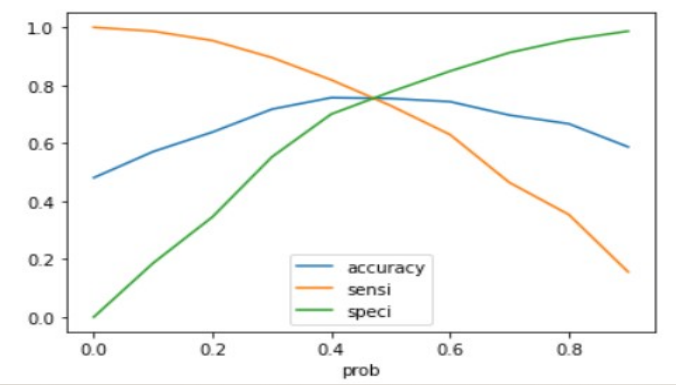


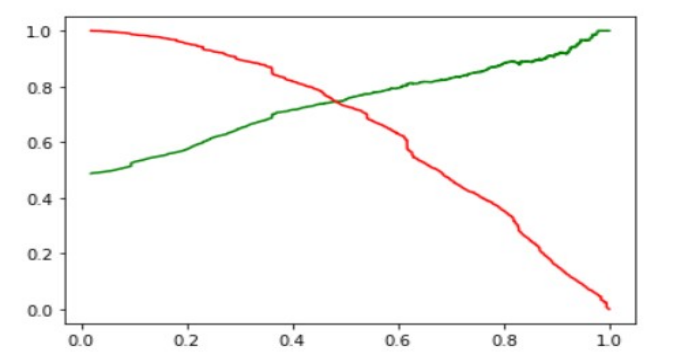
Using Dummy Variables to Get more **statistical** insight

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(TRAINING)





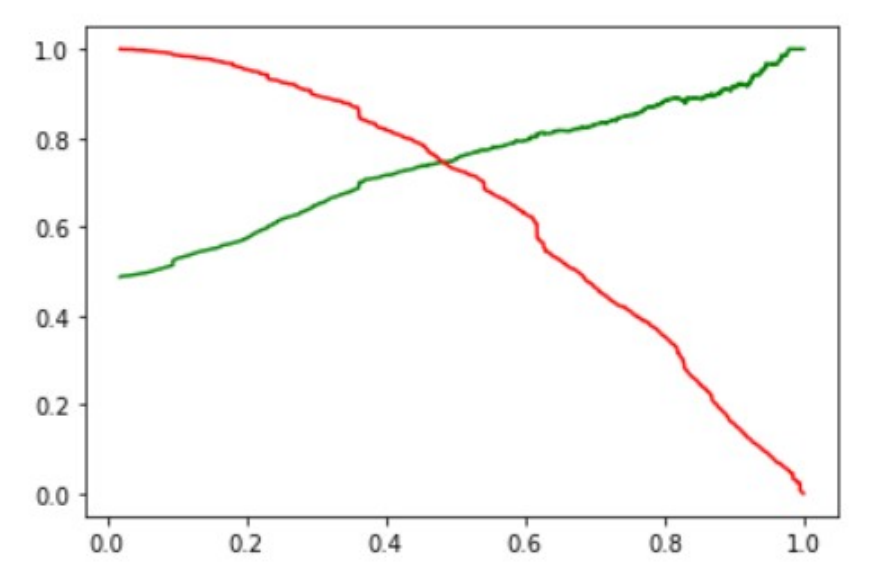
Accuracy , Sensitivity and Specificity

- 84.4% Accuracy
- 84.4% Sensitivity
- 80.5% Specificity

PRECISION AND RECALL

- 70.6% Precision
- 79.5% Recall

MODEL EVALUATION(TESTING)



PRECISION AND RECALL

- 70.6% Precision
- 79.5% Recall

EDA

1. Many leads are generated in the first stage (top), but few of them turn into paying clients in the second stage. To increase lead conversion, you must properly nurture the potential leads during the middle stage (e.g., by educating the leads about the product and maintaining ongoing communication). Sort out the top prospects first from your generated leads. The factors that have the biggest impact on the likelihood that a lead will be converted are "Total Visits," "Total Time Spent on Website," and "Page Views Per Visit."

2. We need you need to keep a list of leads close to hand so you can let them know about new programmes, services, job openings, and upcoming higher education. Keep a close eye on each lead so you can customise the information you provide to them. Carefully present career opportunities, information, or training programmes that best suit the leads' interests. A good strategy for identifying each lead's demands can help you convert leads into customers.

3. Pay attention to leads that have been converted. Engage leads in question-and-answer sessions to gather the pertinent data you require about them. To find out if the leads want to enrol in online courses, make more enquiries and appointments with them.