



LEADS SCORING CASE STUDY

PROBLEM STATEMENT

Problem Statement An online education company, X Education has a low lead conversion rate. To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'. The goal is to build a model which assigns a lead score 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 target is to achieve a lead conversion rate of 80%

Business Objective

- X education wants to know most promising leads.
- For that they want to build a Model which identifies the hot leads.
- Deployment of the model for the future use

SOLUTION METHODOLOGIES

Data cleaning and data manipulation

1. Check and handle duplicate data.
2. Check and handle NA values and missing values.
3. Drop columns, if it contains a large number of missing values and are not useful for the analysis.
4. Imputation of the values, if necessary.
5. Check and handle outliers in data.

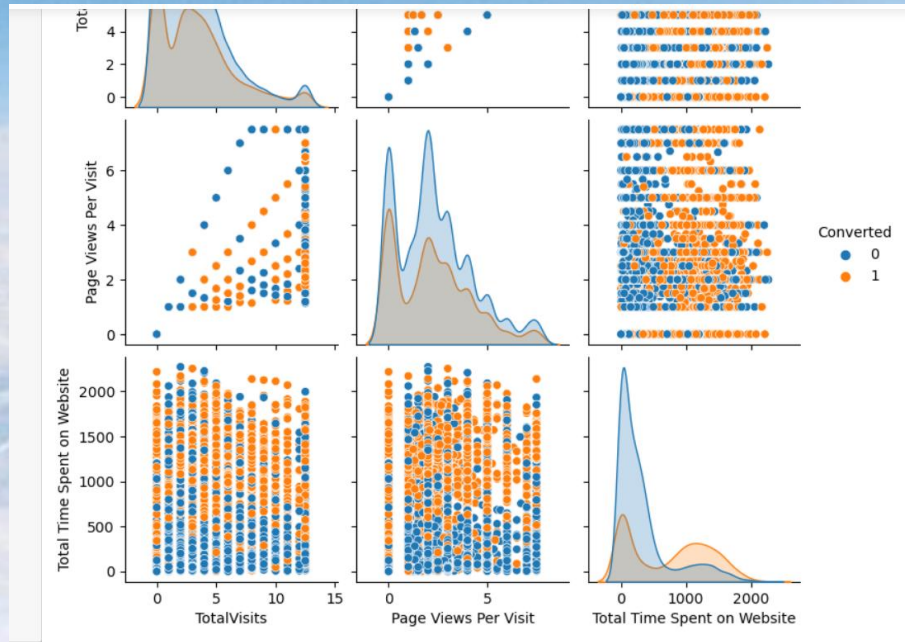
Exploratory Data Analysis (EDA)

1. Univariate data analysis: value count, distribution of variables, etc.
2. Bivariate data analysis: correlation coefficients and pattern between the variables etc.
3. Feature Scaling & Dummy variables and encoding of the data.
4. Classification technique: logistic regression is used for model making and prediction.
5. Validation of the model.
6. Conclusions and recommendations.

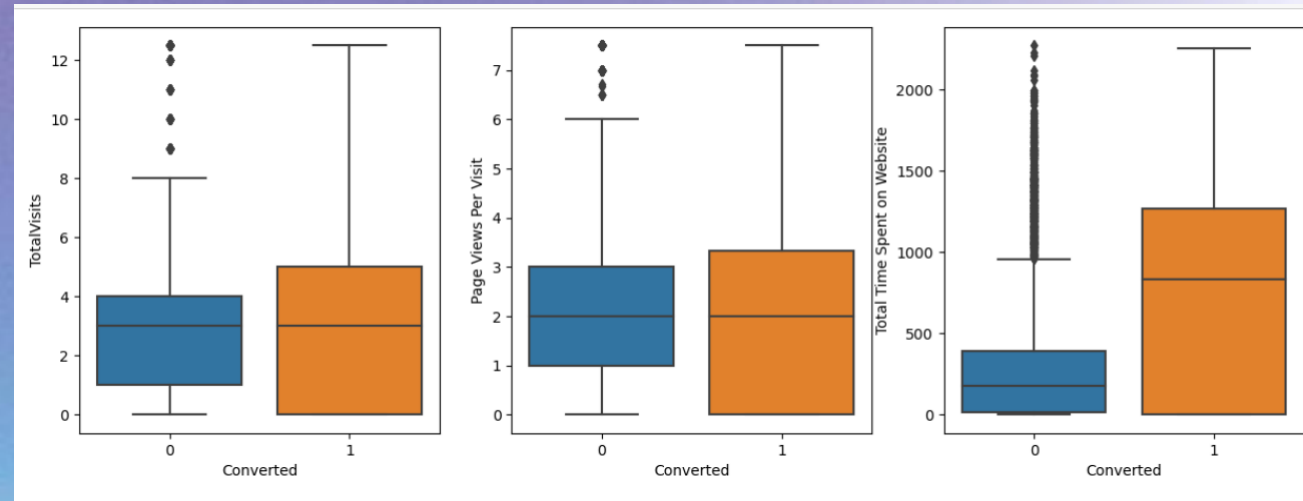
DATA MANIPULATION

- X Removing the “ProspectID” and “Lead Number” which are not necessary for the analysis

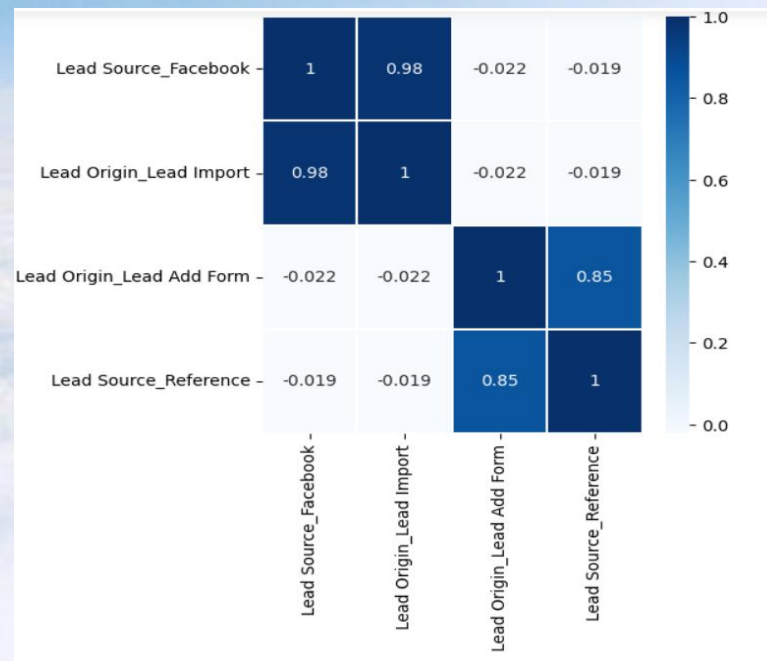
EXPLORATORY DATA ANALYSIS (EDA)



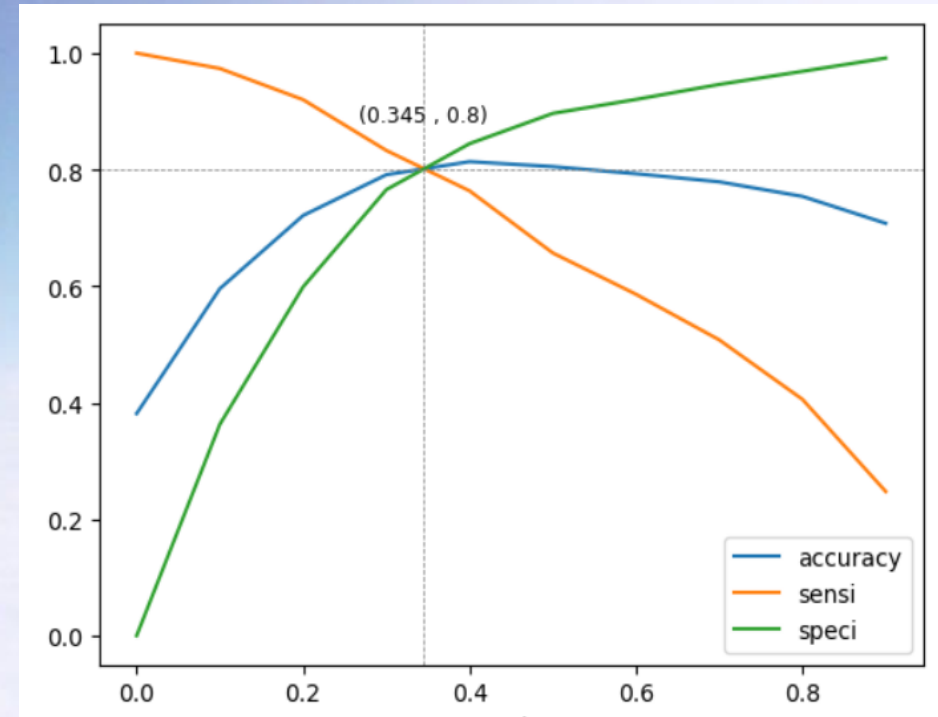
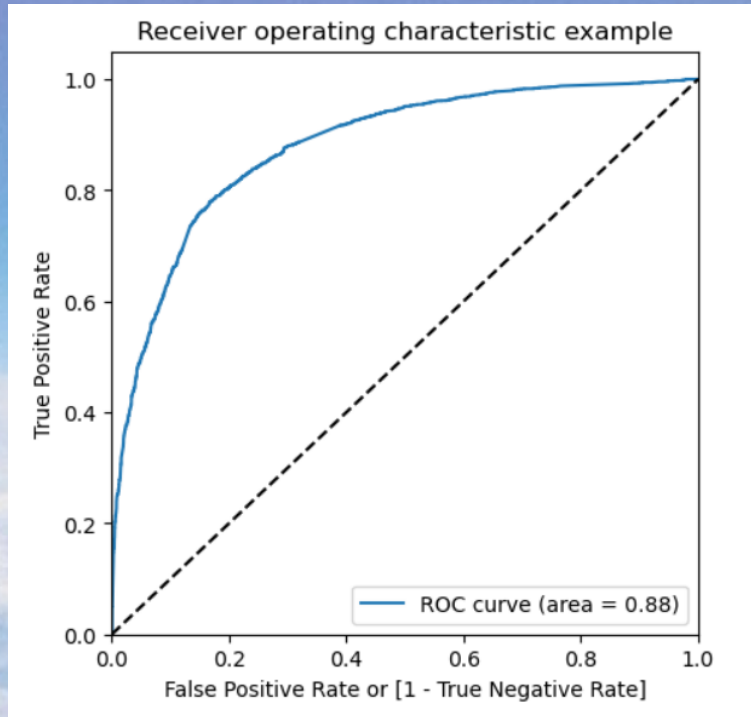
BOX PLOT



HEAT MAP



ROC Curve



CONCLUSION

It was found that the variables that mattered the most in the potential buyers are (In descending order) :

- ▶ The total time spent on the Website.
- ▶ Total number of visits.
- ▶ When the lead source was:
 - Google
 - Direct traffic
 - Organic search
- ▶ When the last activity was:
 - SMS
 - Olark chat conversation
- ▶ When the lead origin is Lead add format.
- ▶ When their current occupation is as a working professional.