Lead Score Case Study

X Education increased lead conversion rate to 80%

Submitted by: Rathina Rajeswari, Ranith Sardar, Praveen Rathi

Submission Date: 18-Jul-2023, Batch: DS C52 23

The problem

Company

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

Context

The company gets a lot of leads, its lead conversion rate is very poor. The company wishes to identify the most potential leads, also known as “Hot Leads”.

Problem statement

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

Challenges deep-dive

Challenge 1

Data handling

To clean the dataset, handle missing values, outliers, scale the numeric features, create dummies for categorical features

Challenge 2

Identify features

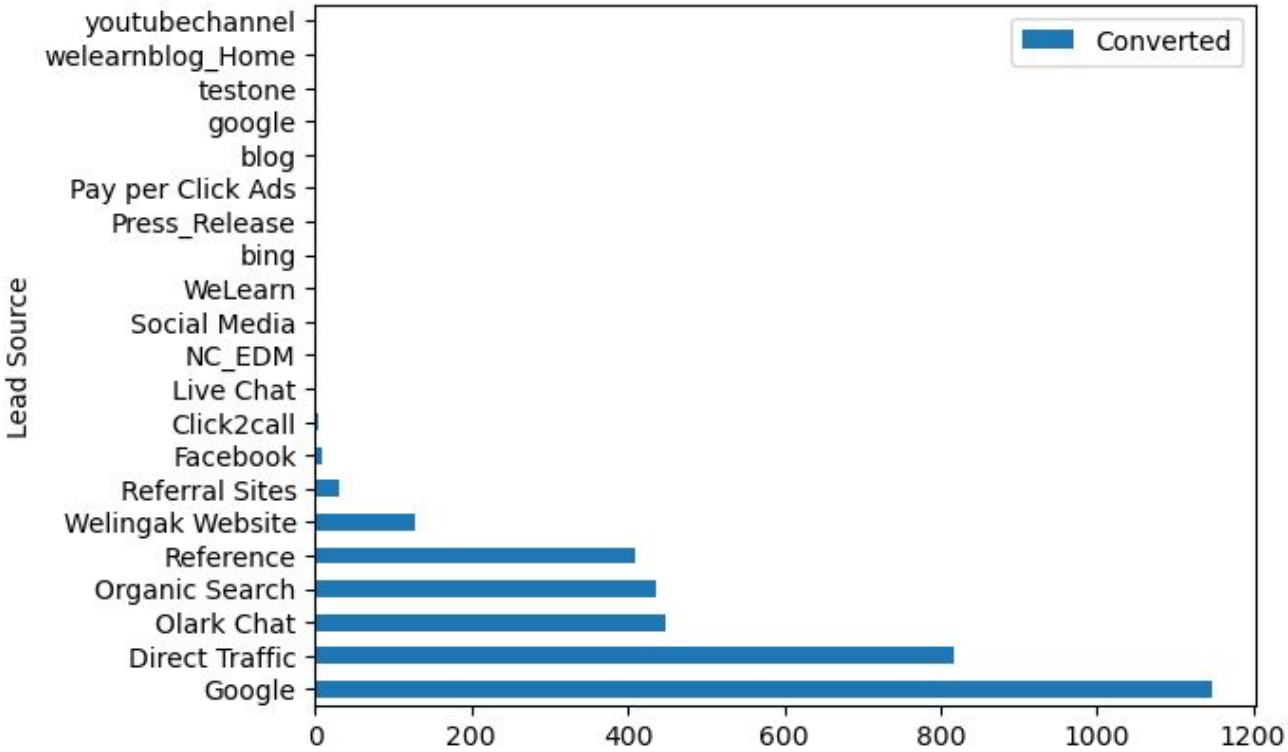
To identify most contributing features through EDA to include in Model which predict very well the higher probability of the leads conversion

Challenge 3

Increase conversion

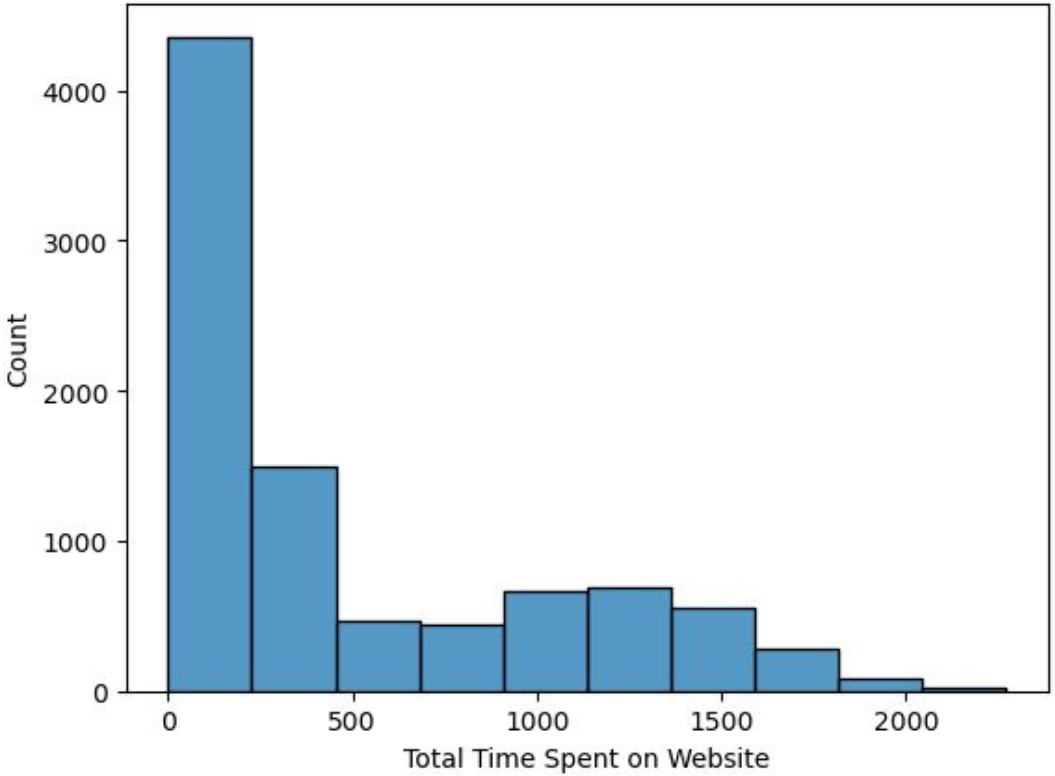
Increase current conversion rate ~39% to ~80% predicting the “Hot leads” through our logistic regression model

EDA - Lead Source



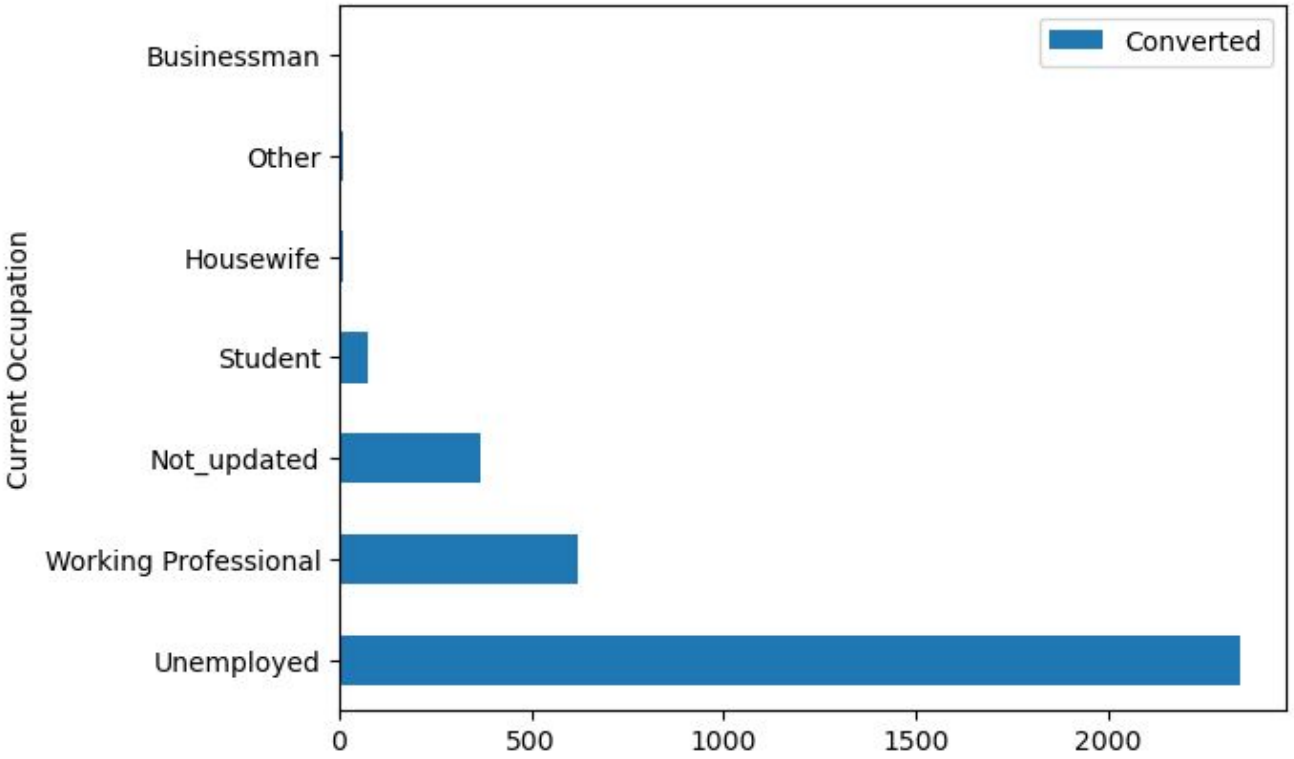
Higher number of leads converted are from Google, Direct , Olark Chat, Organic Search, Reference

EDA - Total Time Spent



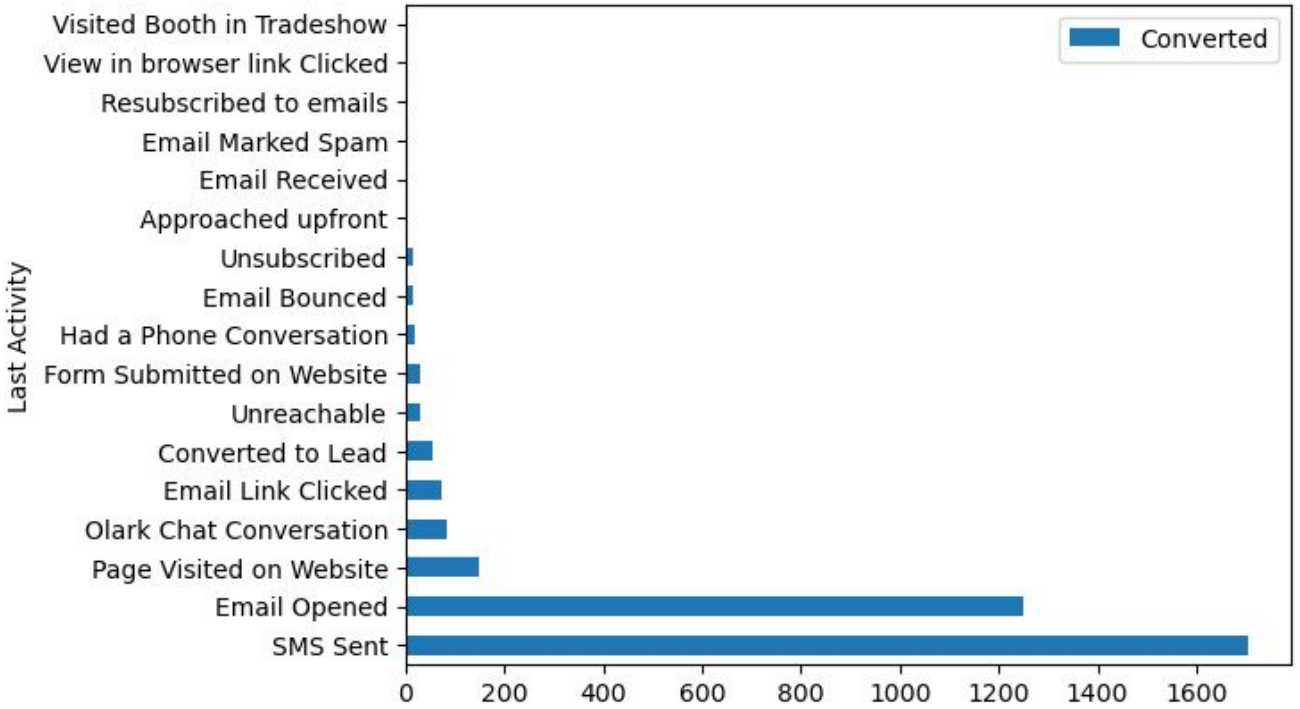
Higher number of leads are spending very less time on website

EDA - Current Occupation



Higher number of leads converted have current occupation as unemployed, working professional

EDA - Last Activity



Higher number of leads converted last activity is SMS sent and Email Opened

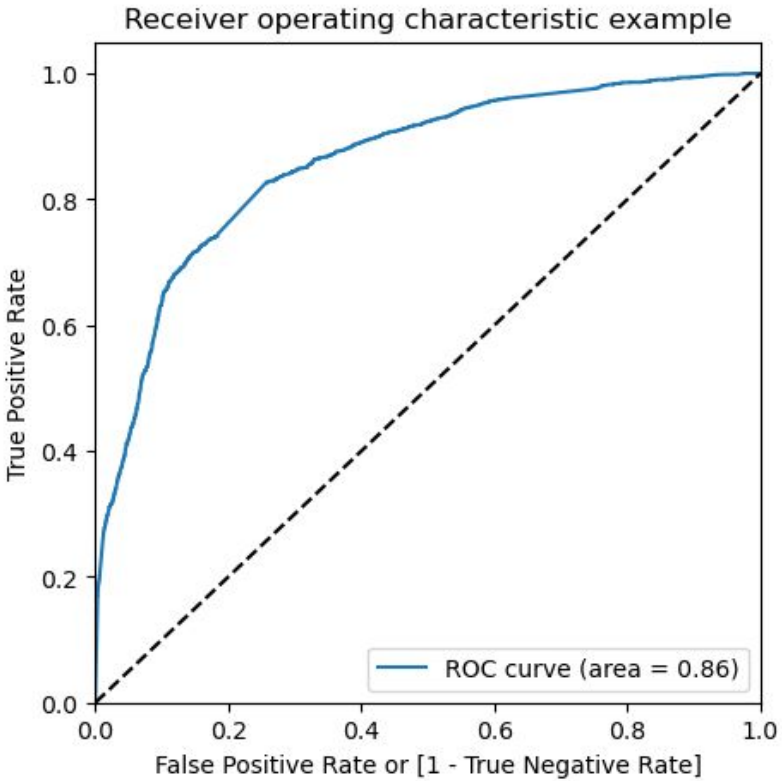
Model Evaluation

Our model is trained with 77.06% accuracy has resulted on test data with 77.52% accuracy

Summary of model

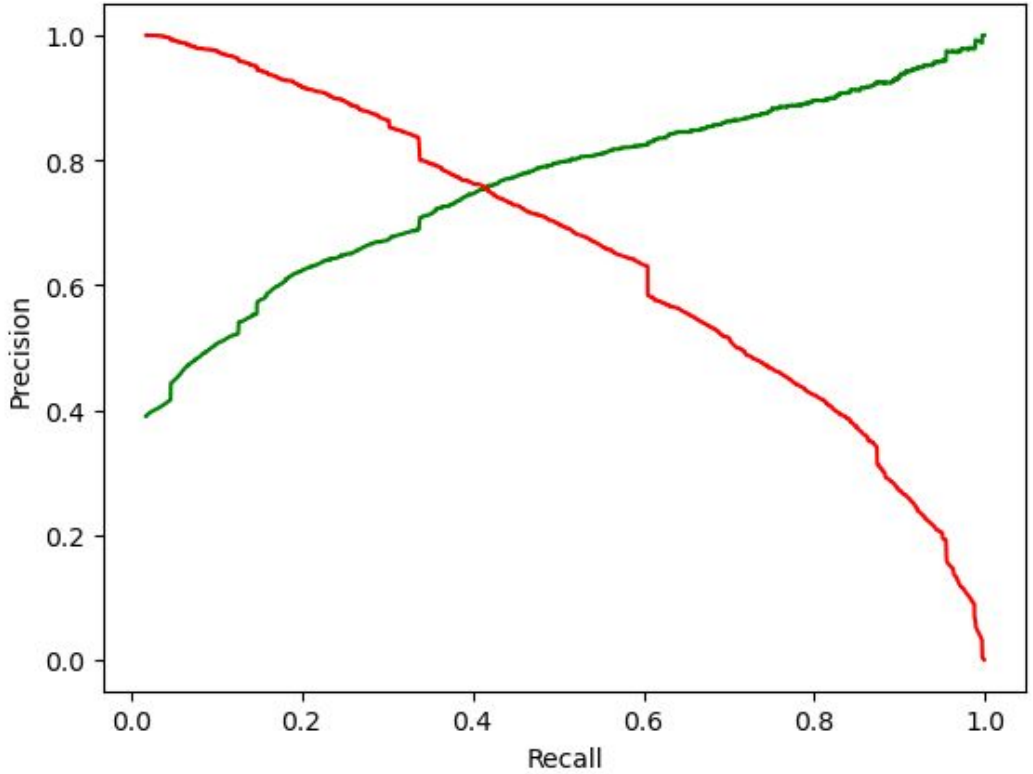
	Train	Test
# Accuracy	77.06%	77.52%
# Sensitivity	64.90%	64.90%
# Specificity	89.80%	89.80%
# Precision	65.67%	66.44%
# Recall	82.89%	83.01%

ROC Curve



Area under ROC curve is 0.86 which is good for this model

Cut off point



Cut off point is trade off between precision & recall i.e. ~0.42 as shown in graph.

Solution

Focus on leads with probability higher than 0.30 to achieve the lead conversion rate to 83%, as predicted by the Model with ~77% accuracy

Most contributing features to higher probability of lead conversion are:

1.) Leads from Google, Direct, Reference, Organic search

2.) Leads currently unemployed or working professional

3.) Leads origin from Add form

Thank you