

Lead Scoring Case Study

SenthilKumar
Sameer Karekar

Objective

To achieve a Lead conversion rate of 80% from the typical 30% by developing Logistic Regression Model

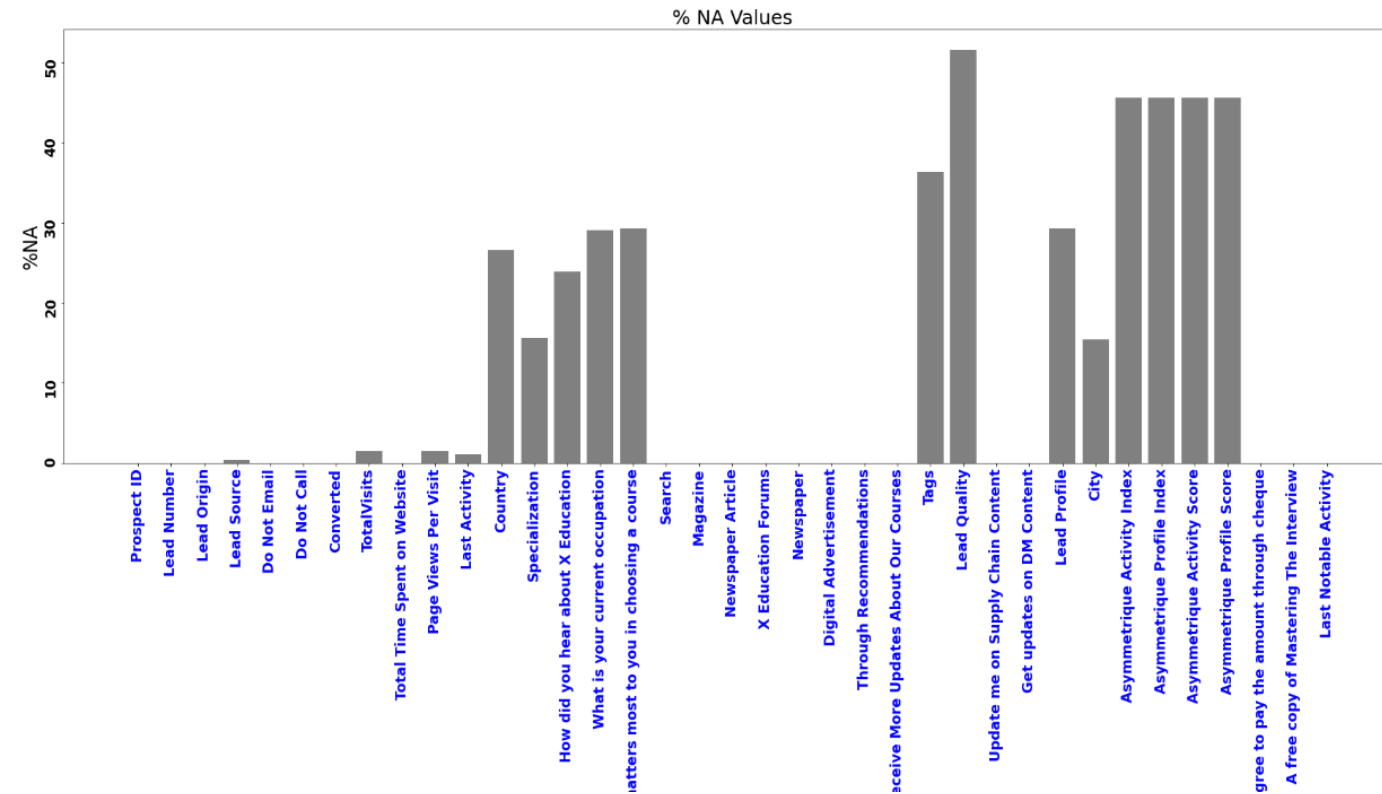
Approach

- Data cleaning
- Data Analysis – EDA
- WoE and IV analysis
- Develop ML model
 - Split Data set
 - Fit Logistic regression model using Train data set
 - Evaluate Model
 - Fit using Test data set
 - Evaluate Model using test data set

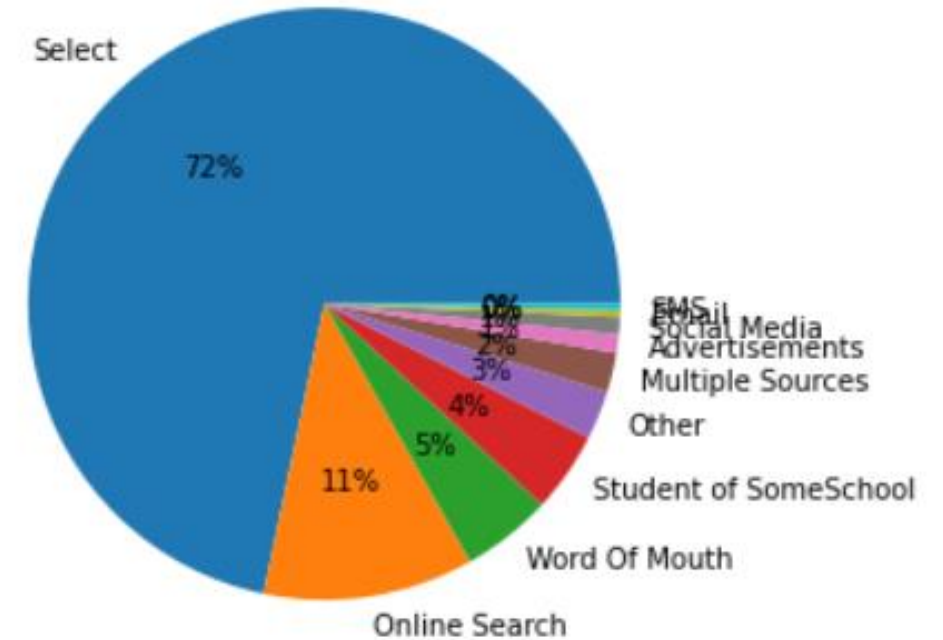
Data Cleaning

- 1) Remove features with Na Values more than 35%
- 2) Remove features containing skewed data

% NaN

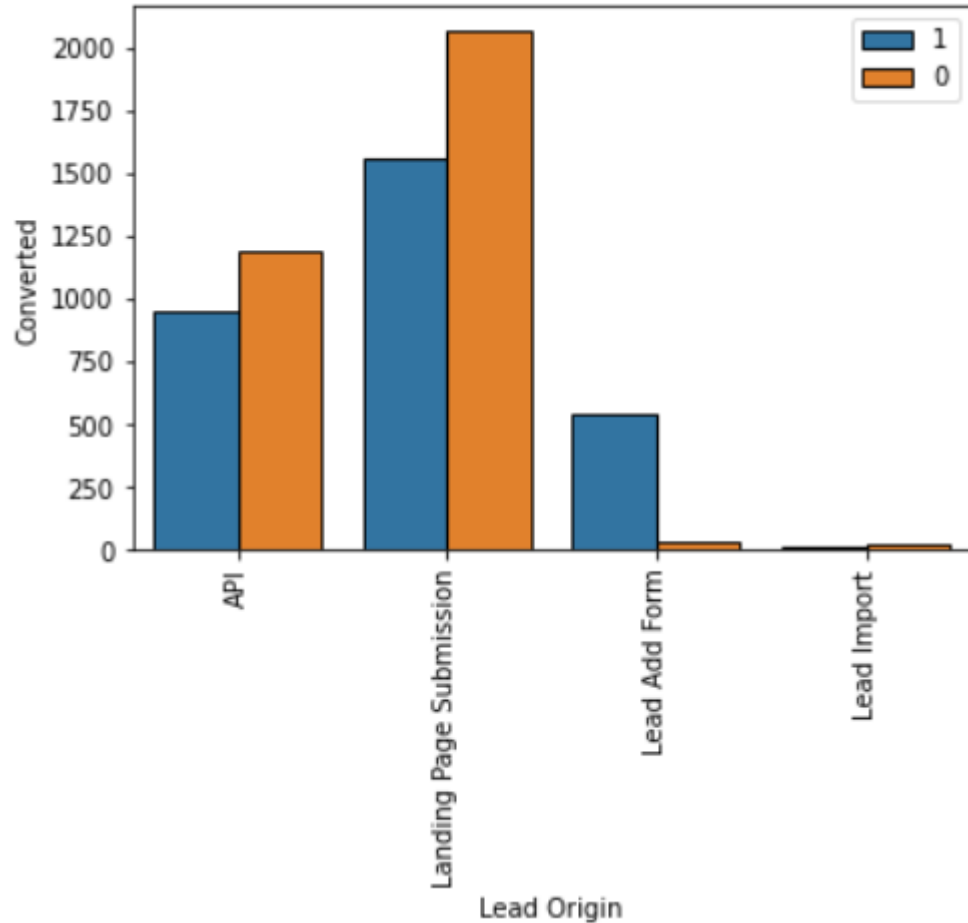


Skewed data

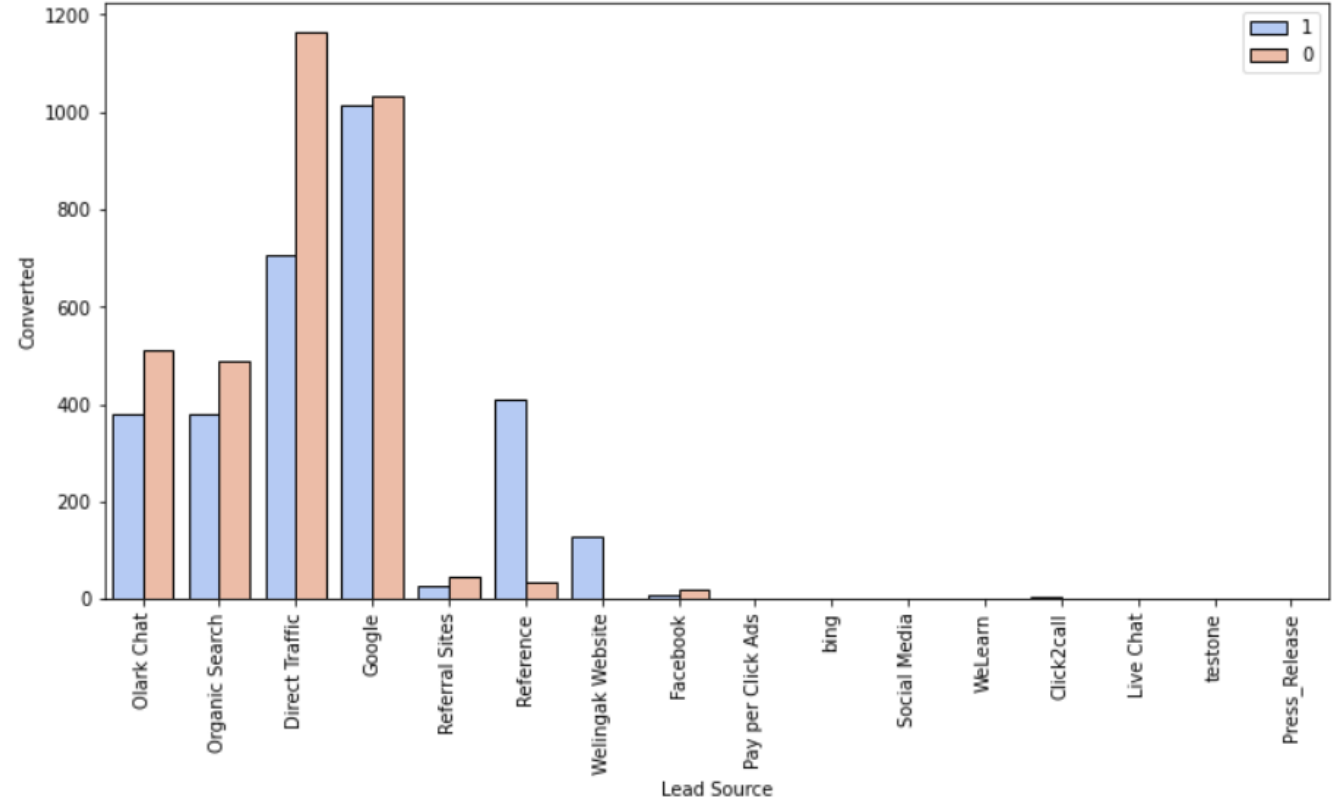


72% of customers did not disclose

EDA – Univariate Analysis (Categorical)

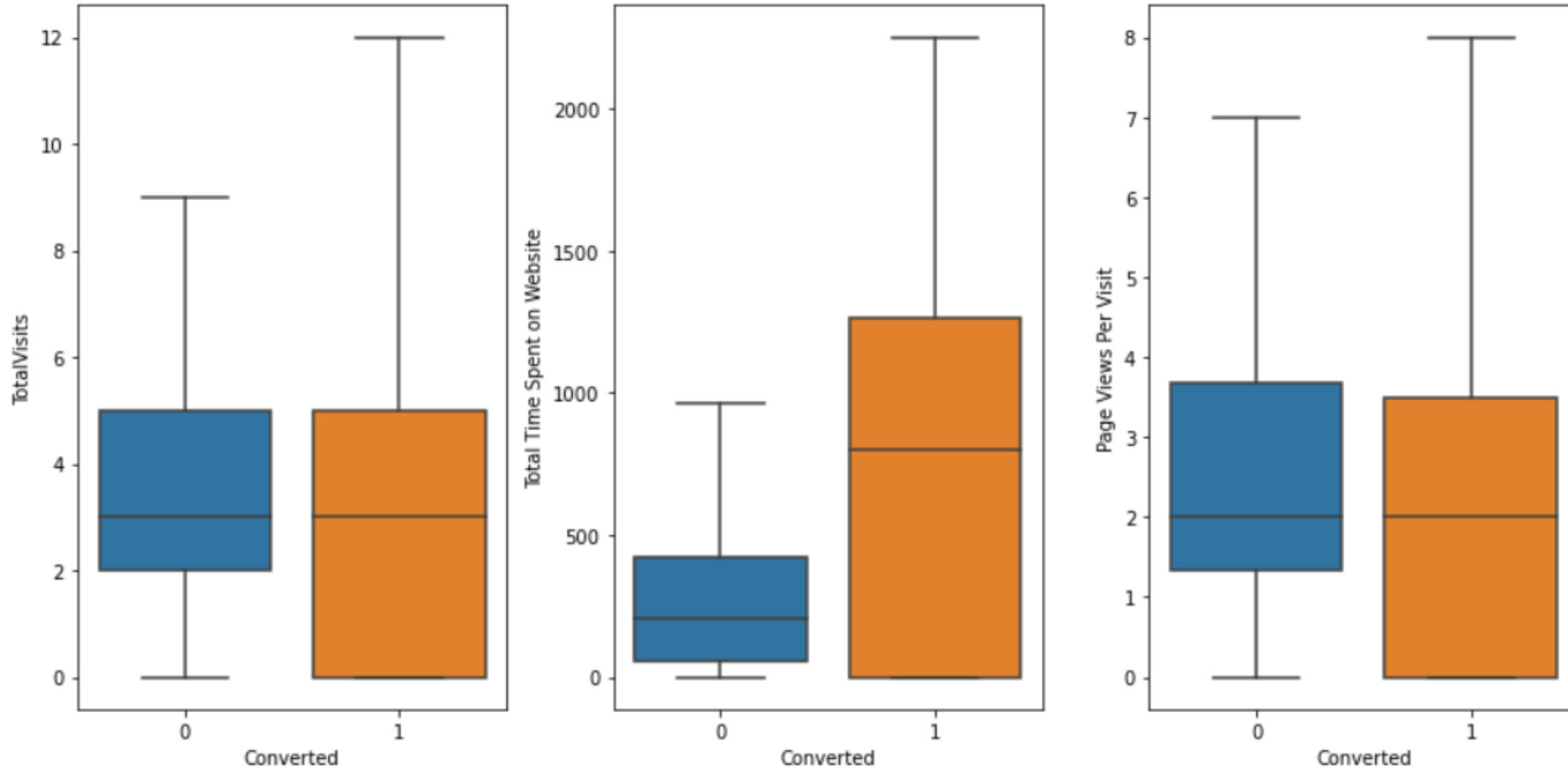


Lead Origin Add Form has higher conversion



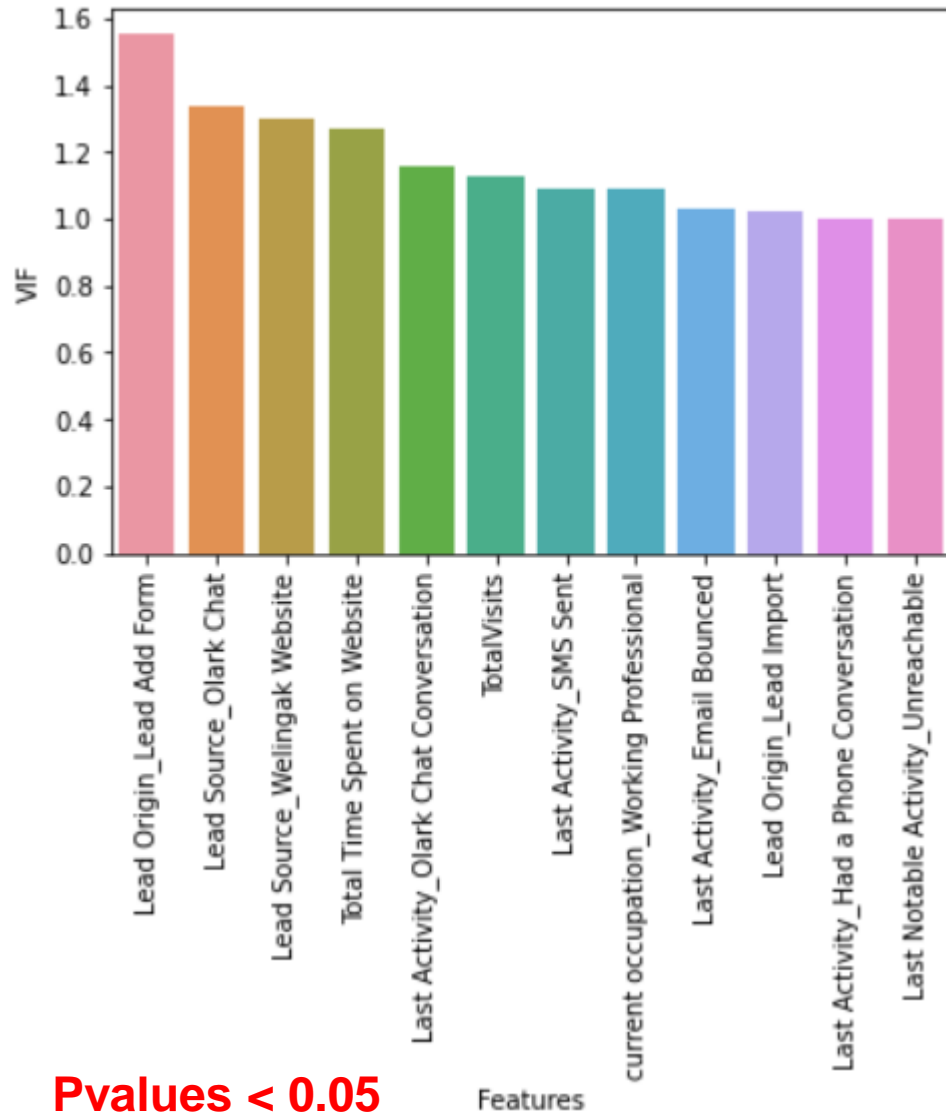
Lead Source Reference and Wellingak website

EDA – Univariate Analysis (Numerical)



Time spent on website has higher conversion

Logistic Regression Model – Train Data

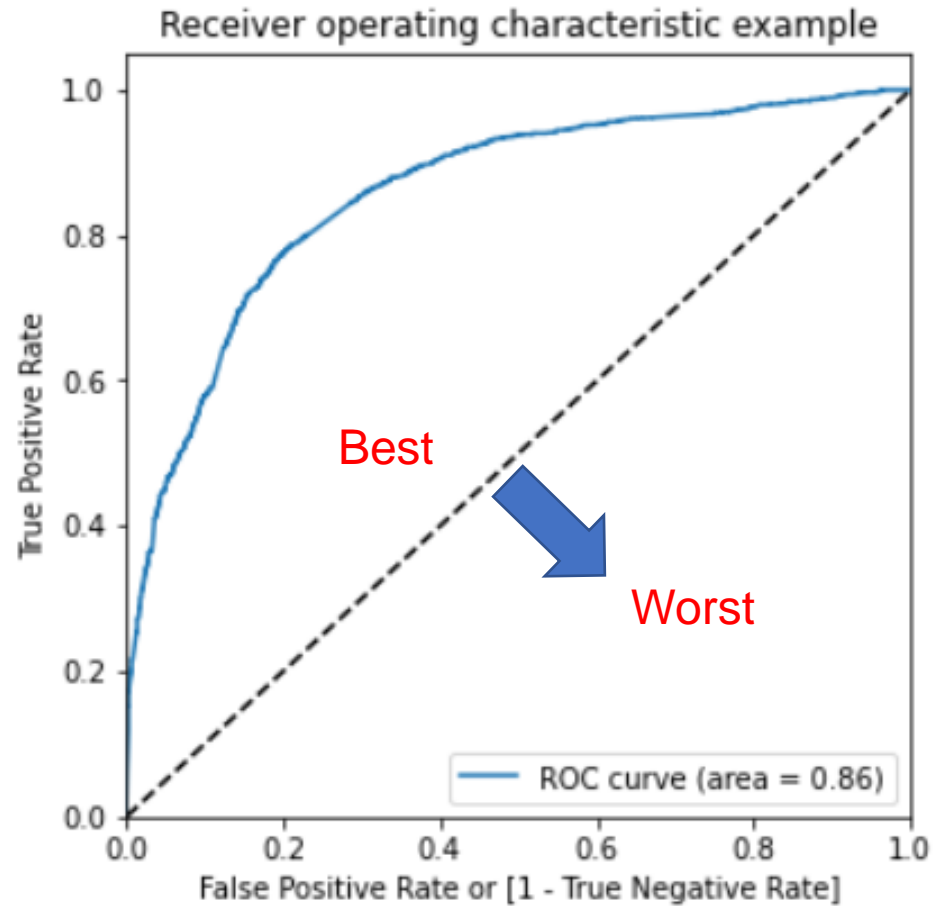


- **Sensitivity or Recall** 0.73
- Specificity 0.83
- False positive rate 0.17
- **Precision** 0.80
- Negative predicted value 0.77

		PREDICTED VALUES	
		Positive (1)	Negative (0)
ACTUAL VALUES	Positive (1)	<div>TP</div> <div>$\text{Sensitivity} = \frac{TP}{(TP + FP)}$$= 1 - \text{Type 2 error}$</div>	<div>FN</div> <div>(Type 2 error with probability = θ)</div>
	Negative (0)	<div>FP</div> <div>(Type 1 error with probability = α)</div>	<div>TN</div> <div>$\text{Specificity} = \frac{TN}{(TN + FP)}$$= 1 - \text{Type 1 error}$</div>

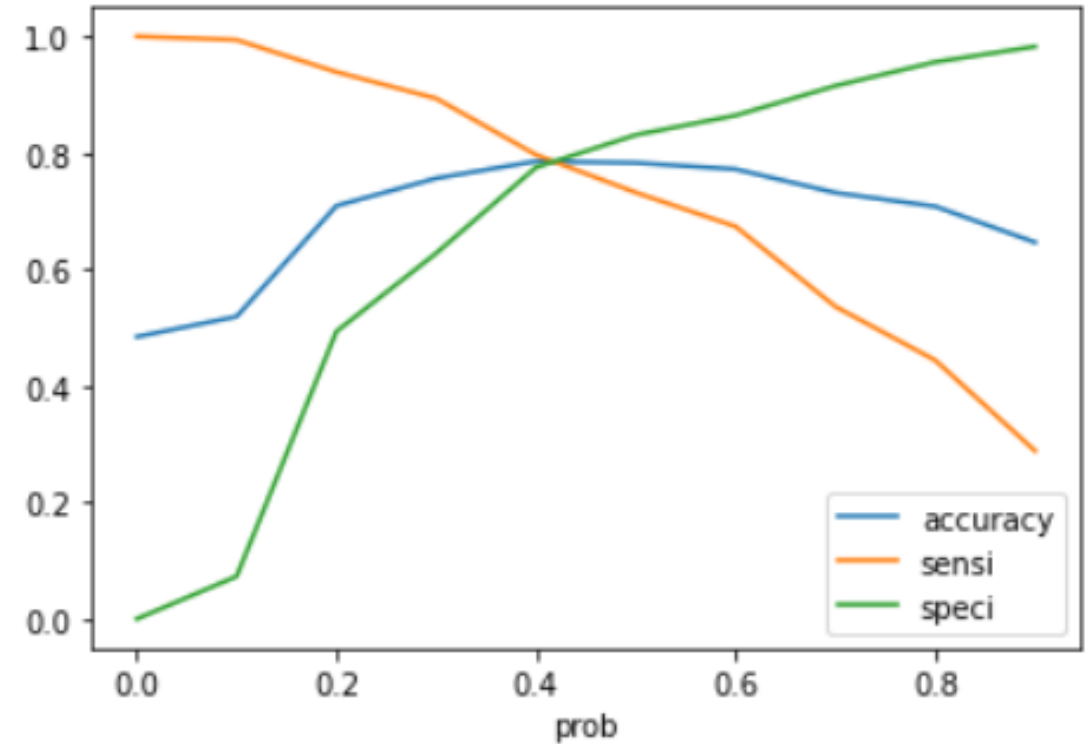
Model precision is 80%

ROC Curve



Model prediction is good

Optimal Cut-off probability



Optimal Cut-off probability is 0.4

Model Evaluation Comparison

Test Data

• Accuracy	0.79
• Sensitivity or Recall	0.8
• Specificity	0.78
• False positive rate	0.22
• Precision	0.77
• Negative predicted value	0.8

Train Data

Accuracy	0.79
Sensitivity or Recall	0.81
Specificity	0.78
False positive rate	0.22
Precision	0.77
Negative predicted value	0.82

Model behaves consistently between train and test data set

Summary

Lead conversion target can be achieved through

- X education portal shall provide more information about courses, benefits, and industry needs.
- Portal shall make customers spend enough time and make frequent visits.
- Carefully review lead origin Add form by seeking more relevant information about customers.
- Target working professionals by adding more industry-relevant courses and customize courses based on needs.
- Olark chat and SMS preferred model of communication
- Focus on Welingak website