

X EDUCATION LEAD SCORING CASE STUDY BY

**MOHD REEYAZ
J.P.RAJKUMAR**

LEAD SCORING CASE STUDY

▶ **Business Objectives:**

- ▶ Increasing the Lead Conversion rate from around 30% to around 80%
- ▶ Current Lead conversion is around 30%.
- ▶ Building the right model to identify and classify the most potential leads tagged as "Hot Leads"
- ▶ The conversion rate from the "Hot Leads" should be around 80%
- ▶ The model should be adjustable to include company's requirement changes.

▶ **Solution Approach:**

- ▶ Data set.
- ▶ Exploratory Data Analysis.
- ▶ Model Building – Logistic Regression.
- ▶ Selection of Optimal Cut-Off Point.
- ▶ Model Evaluation.

LEAD SCORING CASE STUDY

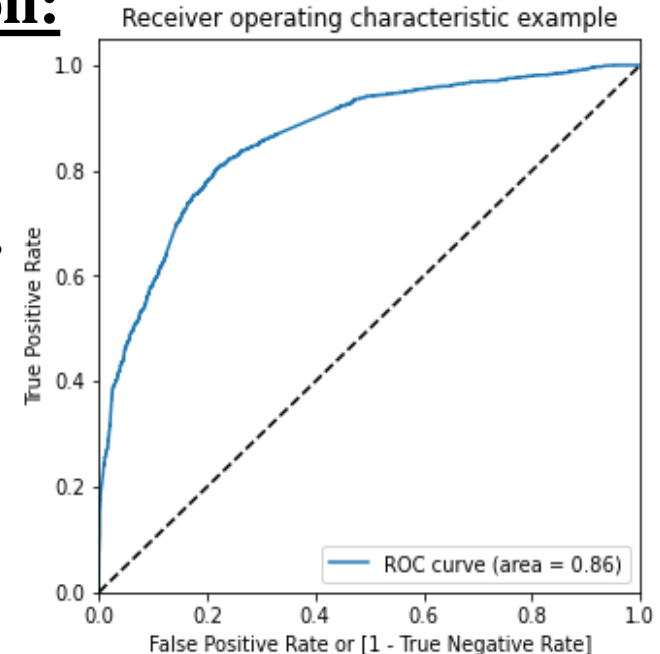
- ▶ **Data Understating and Manipulation:**
- ▶ Lead Data Set provide : 9240 rows and 37 columns.
- ▶ Converting “Select” to “Null” (“Select” implies that user has not selected any value)
- ▶ Drop columns which have only one Unique value, OR the columns which have very less variation.
- ▶ Imputing the “NULL” across different variables, with respective values.
- ▶ Converting categorical variables to dummy features.
- ▶ Test - train split of the data.

LEAD SCORING CASE STUDY

- ▶ **Model Evaluation and Optimization:**

- ▶ **ROC Curves:**

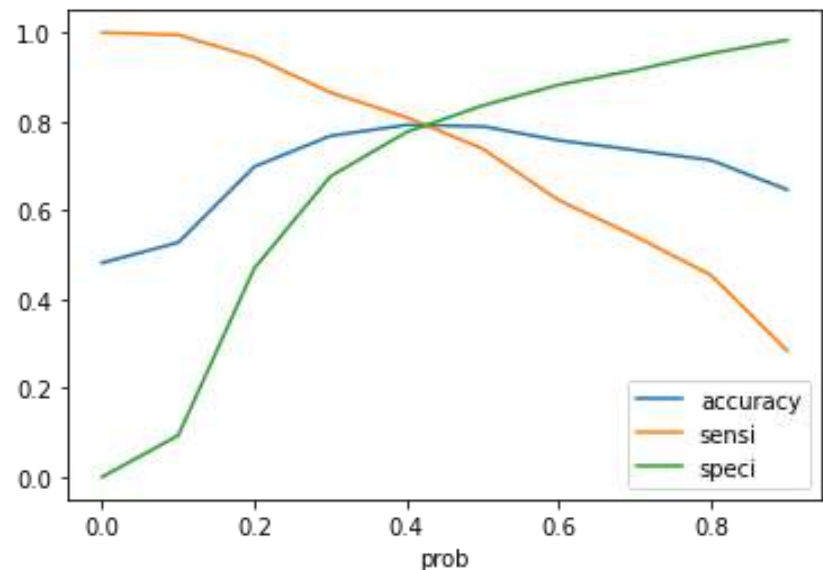
- ▶ Trades off between sensitivity and specificity.
- ▶ For our model, ROC curve is towards the middle,
- ▶ and area under the curve is more as displayed in
- ▶ the fig 1
- ▶ Thus our model is an optimal choice for analysis.



LEAD SCORING CASE STUDY

Finding optimal cut off point:

- ▶ Plotting accuracy, sensitivity and specificity for various probabilities.
- ▶ Cut –Off points is ~ 0.42 Approximately.
- ▶ Where as all the coincide
- ▶ Specifity: ~79%
- ▶ Sensitivity: ~78%
- ▶ Accuracy: ~71%
- ▶ Recall: ~73%
- ▶ Precision: ~80%



LEAD SCORING CASE STUDY

Model Evaluation:

- ▶ Precision and Recall
- ▶ As per business requirement we have selected 0.42 as an cut-off value which gives us better result for accuracy , precision.
- ▶ Accuracy: ~71%
- ▶ Precision: ~80%
- ▶ Recall: ~73%
- ▶ **Top 3 most contributing variables towards the probability of a lead getting converted**

Tags

- CLOSED BY HORIZZON
- LOST TO EINS
- WILL REVERT AFTER READING THE EMAIL

Lead Source

- WELINGAK WEBSITE

Last Notable Activity

- SMS SENT

LEAD SCORING CASE STUDY

- ▶ **Top 3 categorical/dummy variables in the model which should be focused the most to increase the probability of lead conversion**
 - Tags_LOST TO EINS
 - Tags_CLOSED BY HORIZZON
 - Tags_ WILL REVERT AFTER READING THE EMAIL

Conclusion:

- ▶ The model is prepared for prediction of the conversion of the leads. The probability values are generated by the model. The cut-off decided for the model is 0.39. All leads whose probability is generated above this threshold value can be classified as Hot Lead.

LEAD SCORING CASE STUDY

Thank you