Lead Scoring Case Study
(Using Logistic Regression)

03/03/2019

Team Members
Shivali Dalmia
Pranav Kanwar
Viren Agarwal
Biju Kumar

Content

- Problem Statement

- Analysis Methodology

- Business Impact

- Summary

Problem Statement

X Education sells online courses to industry professionals. The company markets its online course , over website and through referral process. Leads driven through 'interest shown' are converted into enrollment.

Currently the company is <u>facing low conversion rate</u> for the leads that have shown interest. By knowing the potential lead – 'Hot Leads' – the sales team can focus on the 'Hot Leads' enabling better conversion rate.

Analysis Methodology
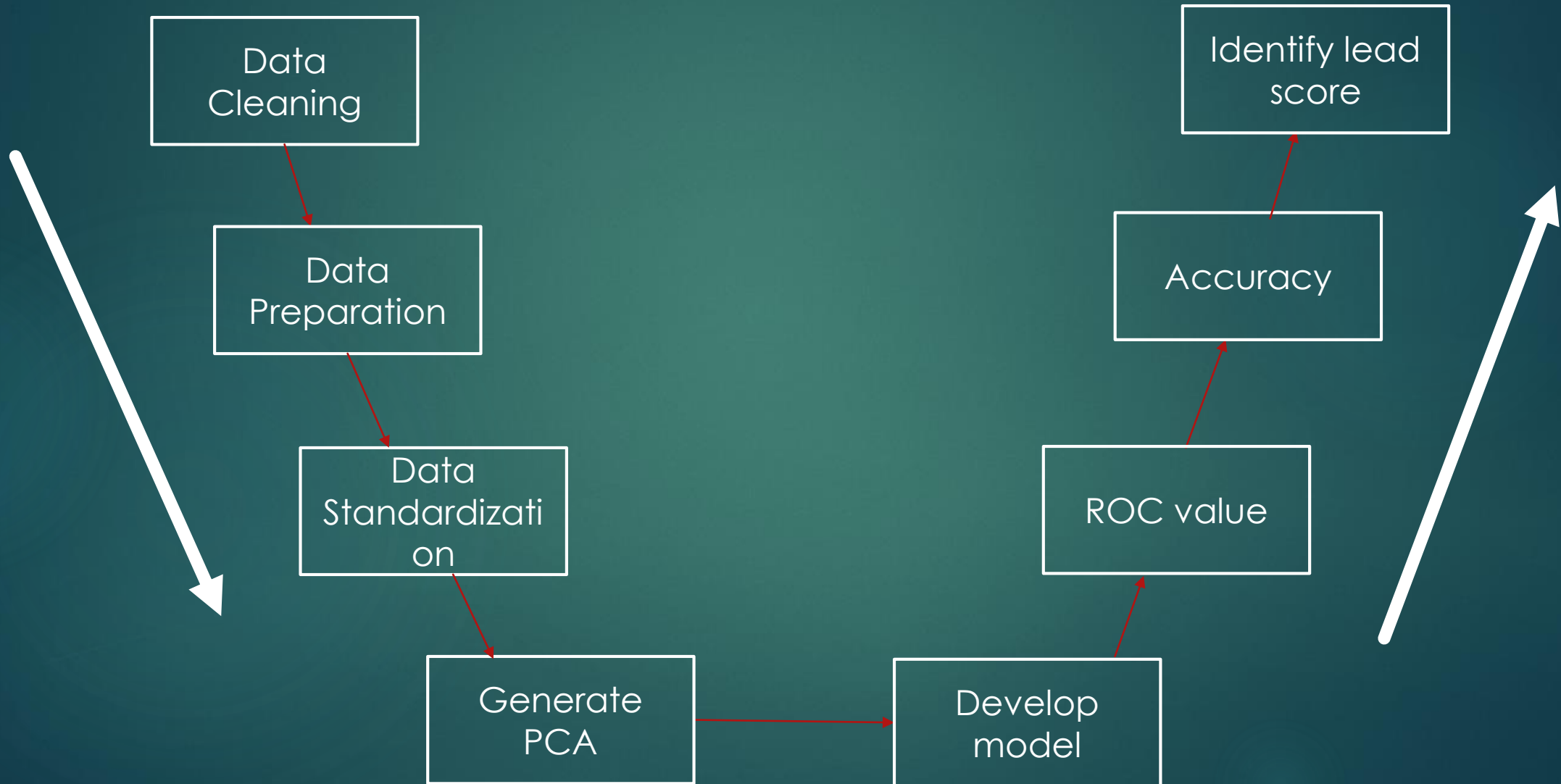
Business Impact (Cutoff Probability)
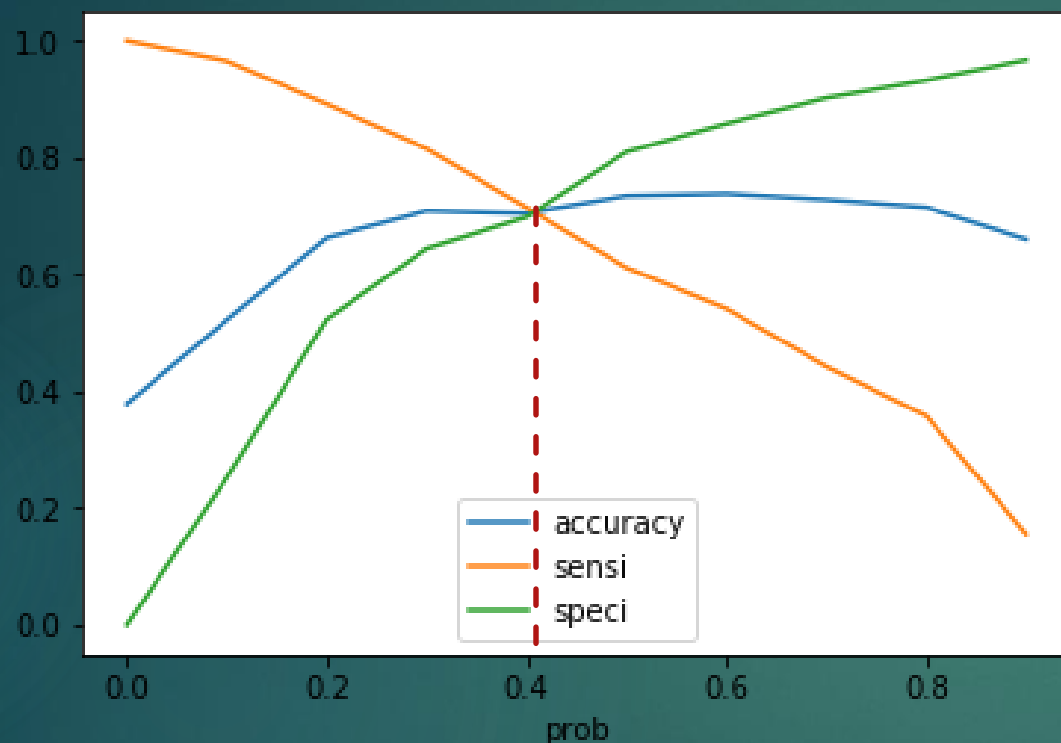
	prob	accuracy	sensi	speci
0.0	0.0	0.378554	1.000000	0.000000
0.1	0.1	0.521821	0.965648	0.251463
0.2	0.2	0.662663	0.891703	0.523142
0.3	0.3	0.709059	0.815429	0.644263
0.4	0.4	0.704430	0.714410	0.698351
0.5	0.5	0.734737	0.611645	0.809718
0.6	0.6	0.738373	0.541485	0.858308
0.7	0.7	0.728124	0.442504	0.902110
0.8	0.8	0.714900	0.357787	0.932435
0.9	0.9	0.660238	0.156041	0.967370

Tightened Cutoff Probability of 0.5 considered in order to drive improvement

Business Impact (Confusion Matrix)

Confusion Matrix
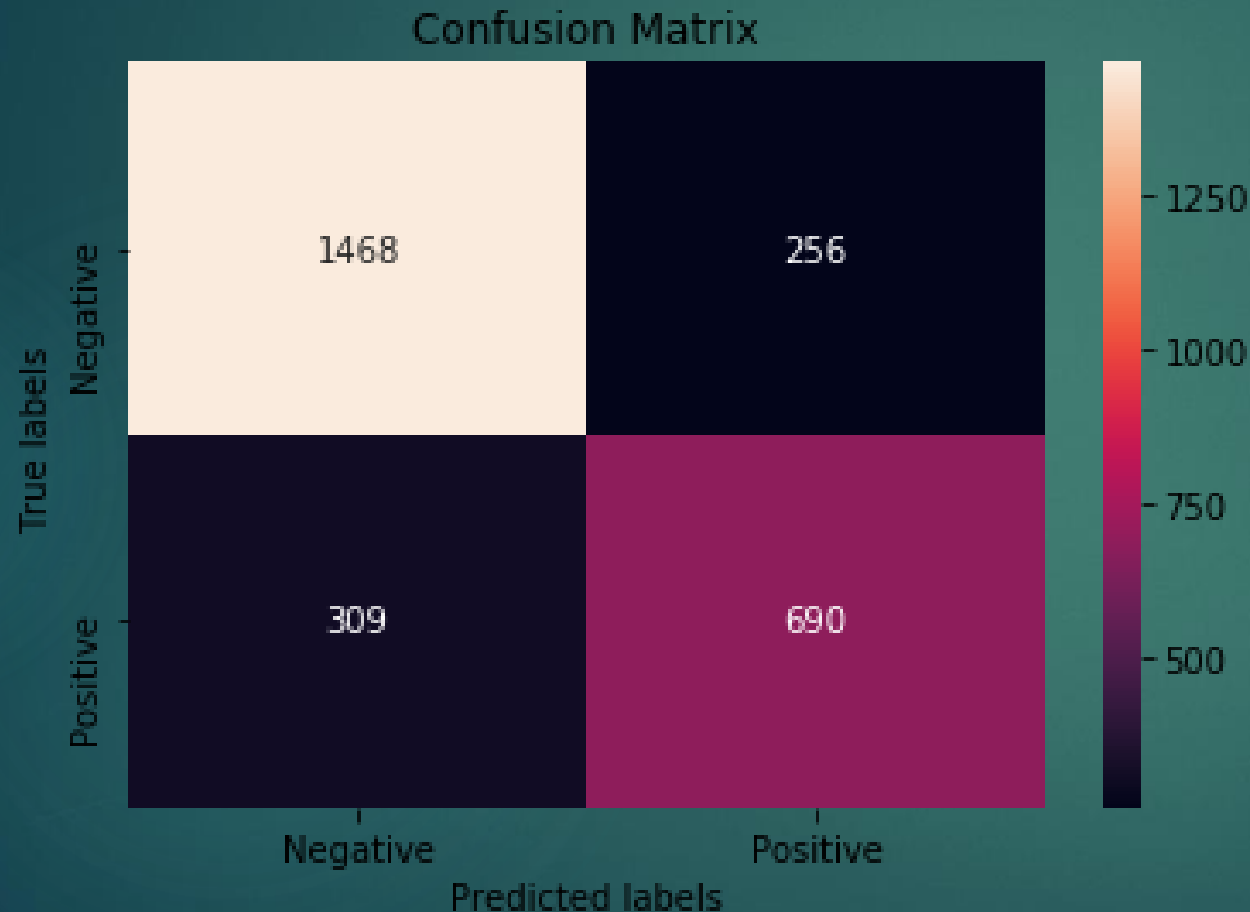
Precision – Probability that predicated 'Yes' is an yes
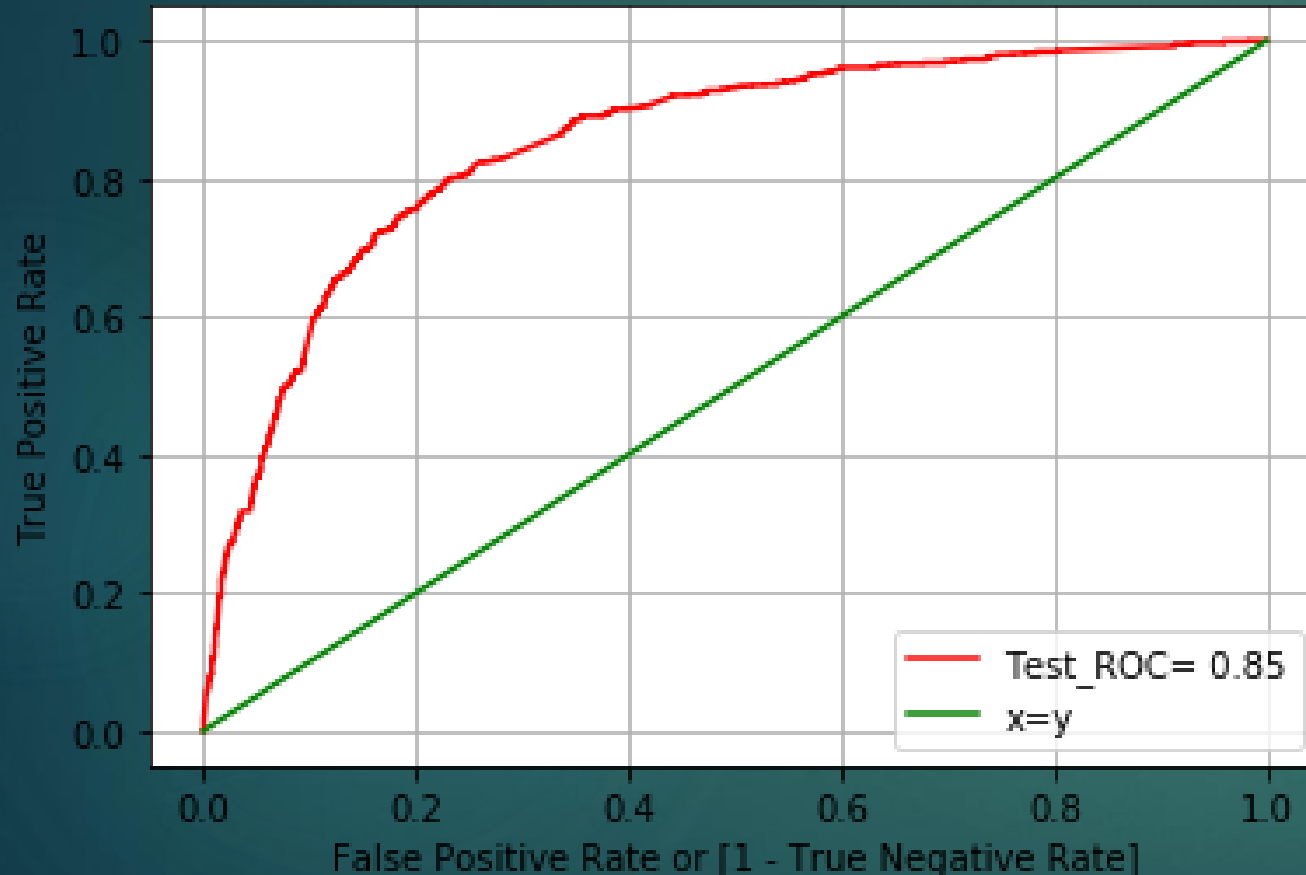
$$Precision = \frac{TP}{TP+FP}$$

Precision = 73%

Recall – Probability that predicated 'Yes' is an yes

$$Recall = \frac{TP}{TP+FN}$$

Recall = 69%

Business Impact (ROC Curve)

Receiver operating characteristic example

ROC Curve (Receiver Operating Characteristic Curve)

It's a tool for predicting the probability of binary outcome

True Positive Rate(TPR) – Sensitivity

$$True\ Positive\ Rate\ (TPR) = \frac{True\ Positives}{True\ Positives + False\ Negatives} = \frac{TP}{TP + FN}$$

False Positive Rate(TPR) – Specificity

$$False\ Positive\ Rate\ (FPR) = \frac{False\ Positives}{True\ Negatives + False\ Positives} = \frac{FP}{TN + FP}$$

Summary

- Sales team could use the lead score to target potential leads.
- Reinforce communication channel, to have better conversion of the lead .
- Current cutoff score is 50%, but could reduce this further to 40%, to improve the coverage.

index		Prospect ID	Lead Number	Lead Score	Converted
0	0	7927b2df-8bba-4d29-b9a2-b6e0beafe620	660737	27.0	0
1	1	2a272436-5132-4136-86fa-dcc88c88f482	660728	12.0	0
2	2	8cc8c611-a219-4f35-ad23-fdfd2656bd8a	660727	63.0	1
3	3	0cc2df48-7cf4-4e39-9de9-19797f9b38cc	660719	9.0	0
4	4	3256f628-e534-4826-9d63-4a8b88782852	660681	35.0	1

Potential Target

Thank you