Lead Case Study

Name:

1. Navneet Sharma

Abstract

X Education Institute have some business problem. Institute wants to an application model which can make prediction about hot lead to convert in customer. X education wants to help them to find most promising lead.

In problem statement, Institute needs top factors to focus on to get more conversion.

This is classification type issue. We will use Classification logistic regression model. We are going to assign lead score 0 to 100 on base of probability. Education has target lead conversion rate to be around

Problem solving methodology

We have inspected all data features and annualize that data problem is classification problem.

Used Logistic regression for solving classification problem.

Data size is more than 90, 000 rows and columns 37.

Convert all variable in dummy formate

Used REF to select top featured variables.

Make two model

1. First model have cut-off point > 0.5
2. Second model have cut-off point >0.3

Lead case Analysis

Create first model “model1” with greater that 0.5 cutoff.

Model accuracy : 0.92

Sensitivity: 0.94

Specificity: 0.94

```
#Specificity
TN/(TN+FP)
```

```
0.948271446862996
```

```
## Check overall accuracy
print(metrics.accuracy_score(df_pred.Converted , df_pred.model_Prediction))
```

```
0.920642418516769
```

```
# Sensitivity
TP/(TP+FN)
```

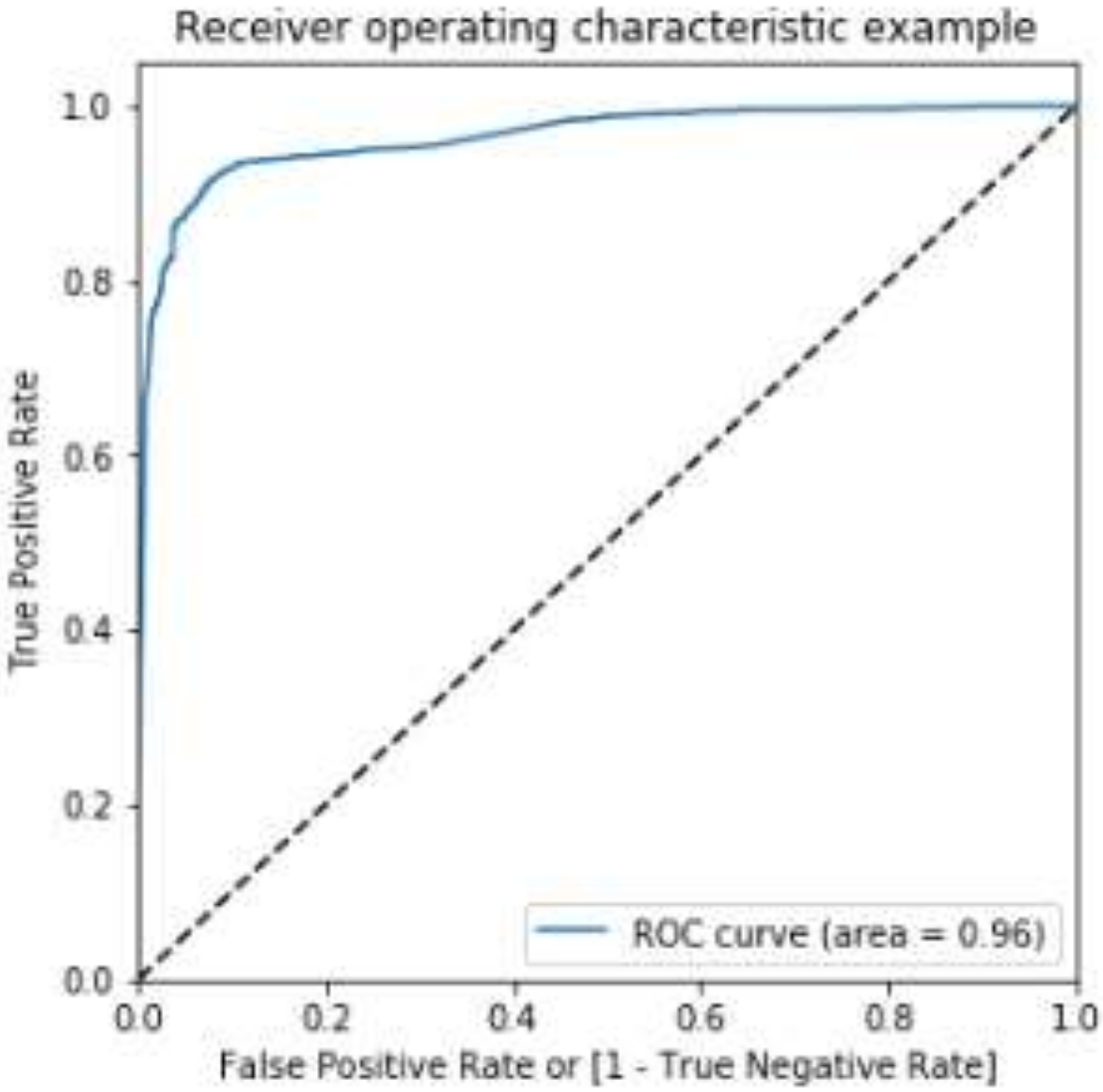
```
0.876533115290269
```

Lead case Analysis Cont..

ROC Curve:

Relation between True Positive rate and False Positive rate or (1- True Negative Rate)

ROC covering 0.96 area.

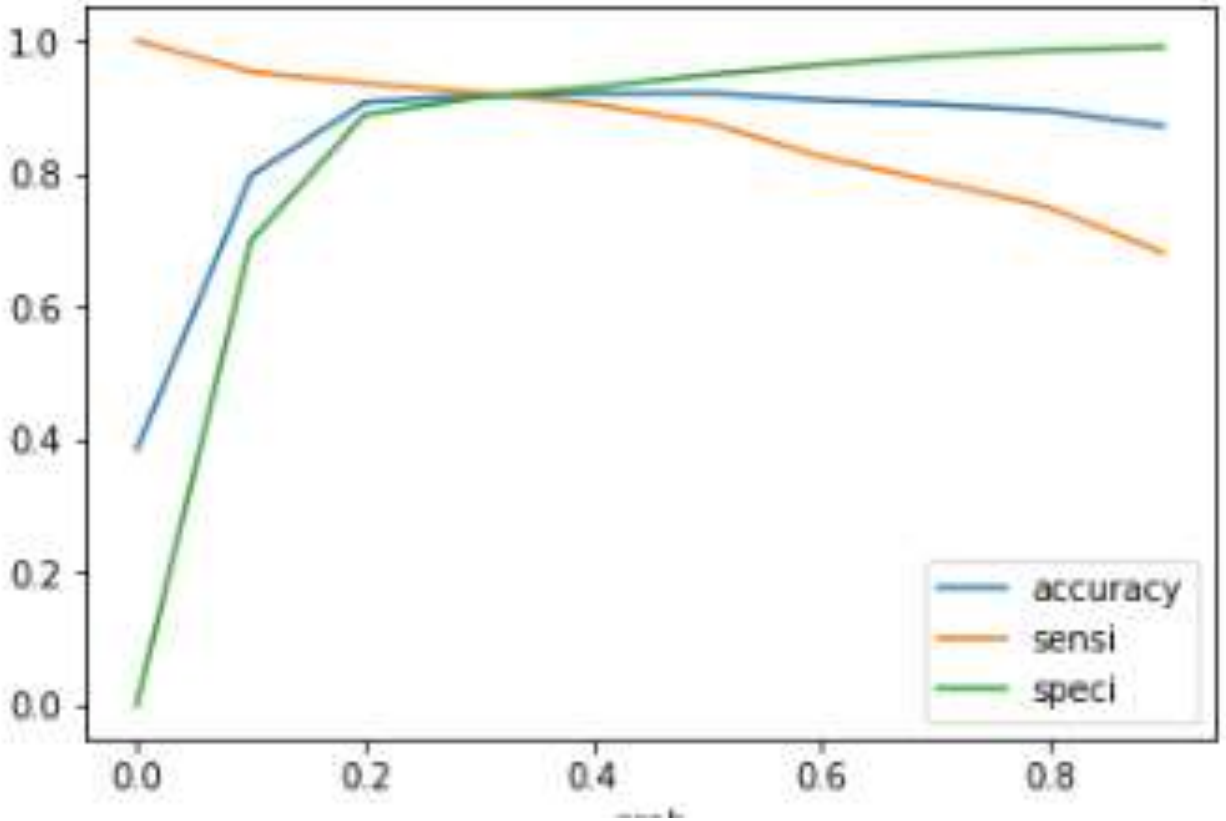


Lead case Analysis Cont. .

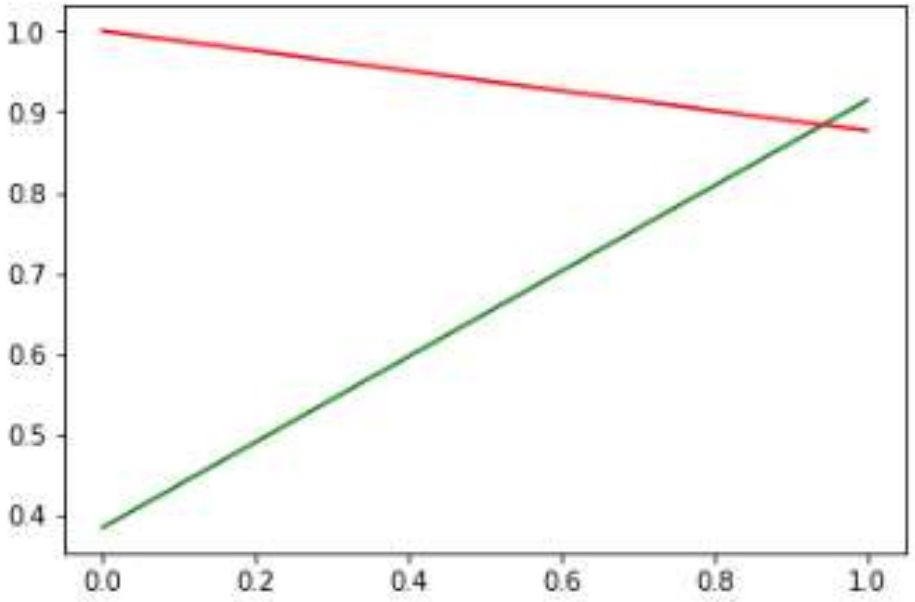
Optimal Cut-off.

Find best optimal cut-off on which Sensitivity and specificity would be balanced.

We found that 0.3 is best fit cut-off.



Precision Recall Score



Precision Score : 0.91

Recall Score : 0.87

Lead Scoring result

	Converted	Converted_prob	ID	model_Prediction	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted	Lead_Score
0	0	0.171861	3009	0	1	1	0	0	0	0	0	0	0	0	0	17
1	0	0.119526	1012	0	1	1	0	0	0	0	0	0	0	0	0	12
2	0	0.000422	9226	0	1	0	0	0	0	0	0	0	0	0	0	0
3	1	0.938607	4750	1	1	1	1	1	1	1	1	1	1	1	1	94
4	1	0.995664	7987	1	1	1	1	1	1	1	1	1	1	1	1	100

Assigned lead score to all customer on the base of chance to conversion.

High score high probability to take lead hot to conversion.

Thank You