



# Lead Score Case Study

---

By Indranil Banerjee





## Problem Statement

---

- The education organization want to increase the leading conversation rate from 30 percentage to 80 percentage To create more customer from the given prospects .



# Solution Steps

---

Read the data

Do descriptive analysis

Do basic inference on the data (deleting the features which has been provided by the marketing team, deleting the features where number of data is less, combining categorical values, deleting categorical features which has high skew)

Visualise the data find the relation in between features and their effect on the target variable

Binning of numerical variable to categorical variable if the uniqueness is very less

Do outlier treatment through capping

Create dummies of the categorical feature

Model building basis steps (splitting the data, scaling the data)

Feature selection through (RFE, p-value, VIF)

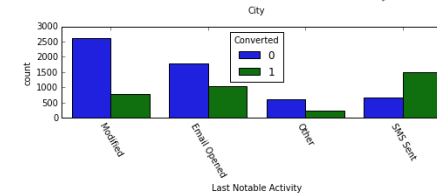
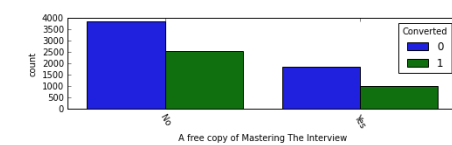
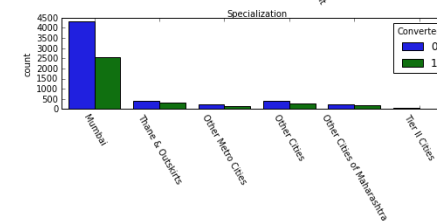
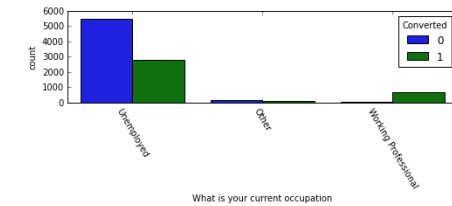
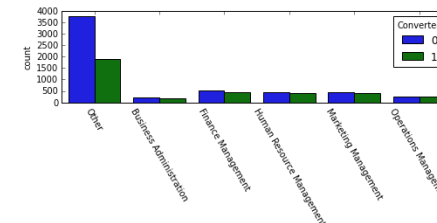
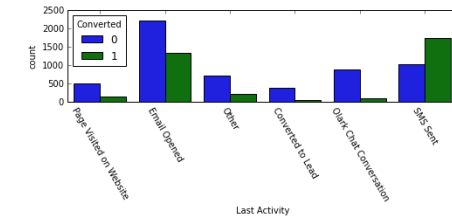
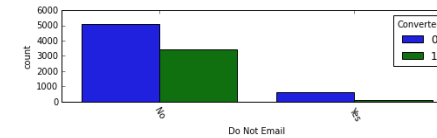
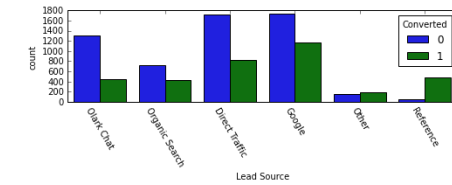
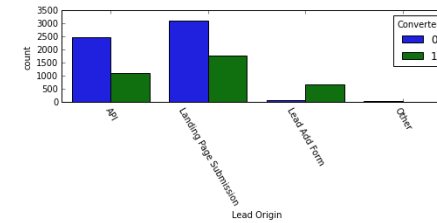
Generate balance model by find the best cut off probability between accuracy, sensitivity, specificity.

See the model performance on test make necessary tweak to probability cut off so that we get the desired cut-off of above 80%.

# Visualization Findings

- last notable activity is SMS sent then it is high likely the student will be going to convert if modified then it is less likely
- for mumbai location most has converted and not converted as well.
- working professional are most like to convert
- for specialization it is 50-50 chance of conversion(business administration, finance management, human resource management, marketing management, operation management)
- if the last activity is sending SMS high likely to convert ,olkar chat conversation ,visited website page less likely to convert
- if the lead source is olkar chat less chance to convert ,for reference convert chance is very high
- if the lead origin is adding form then it high chance of convert

categorical values analysis with Converted



# Recommendation

## Short Term

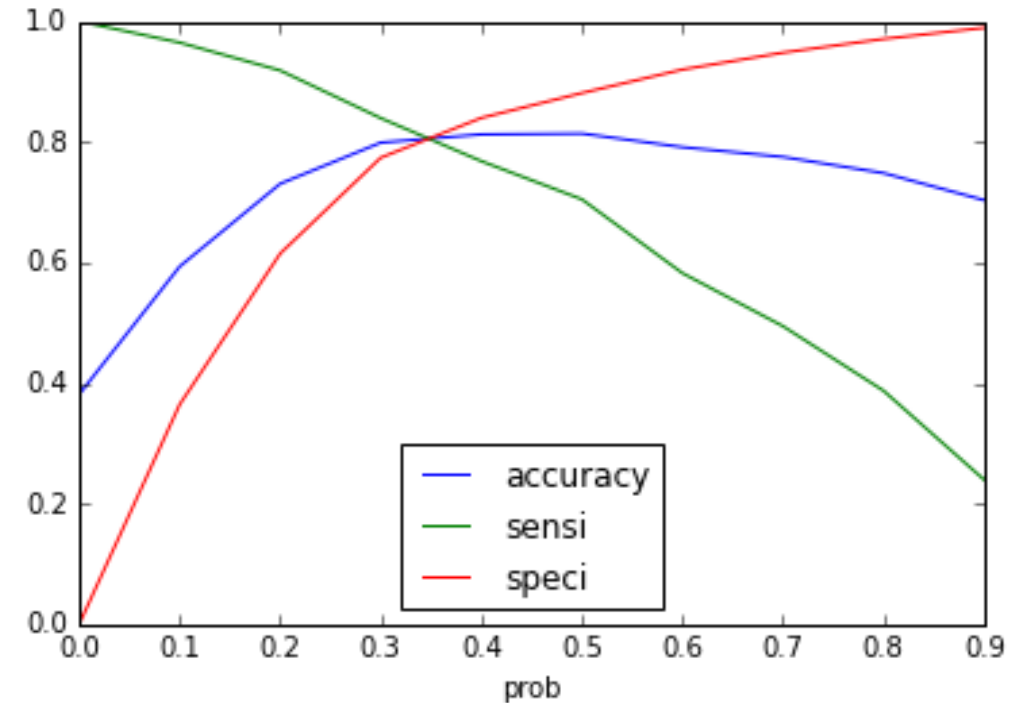
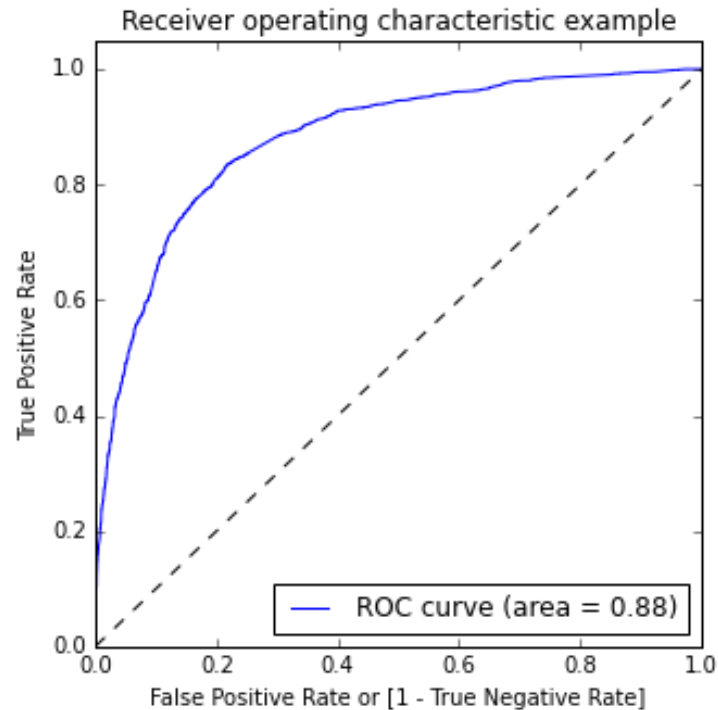
- should call to more working professional
- call to the customer who are coming through reference
- call the who are willing to get the emails
- concentrate on the customer whose lead has been originated from Lead Import, Quick Add Form
- The customer who latest activity is sent SMS concentrate on that customer

## Long Term

- any lead which in coming from olkar chat ,if the last activity is also Olkar chat ,if someone is visiting the website they are less likely to convert try to make our own chat system more customer friendly and attractive so that user come organically ,rather than calling them.

# Conclusion

- As Pre CEO has asked that the lead conversion should happened more than 80% we have achieve it through sensitivity more than 80% by making the cut of probability less than .35 meaning 3 to increase the sensitivity.
- The area under the ROC curve is more than 88% which is very good that means the overfitting is very less



# Thanks You

