

# LEAD SCORING CASE STUDY

## SUBMISSION

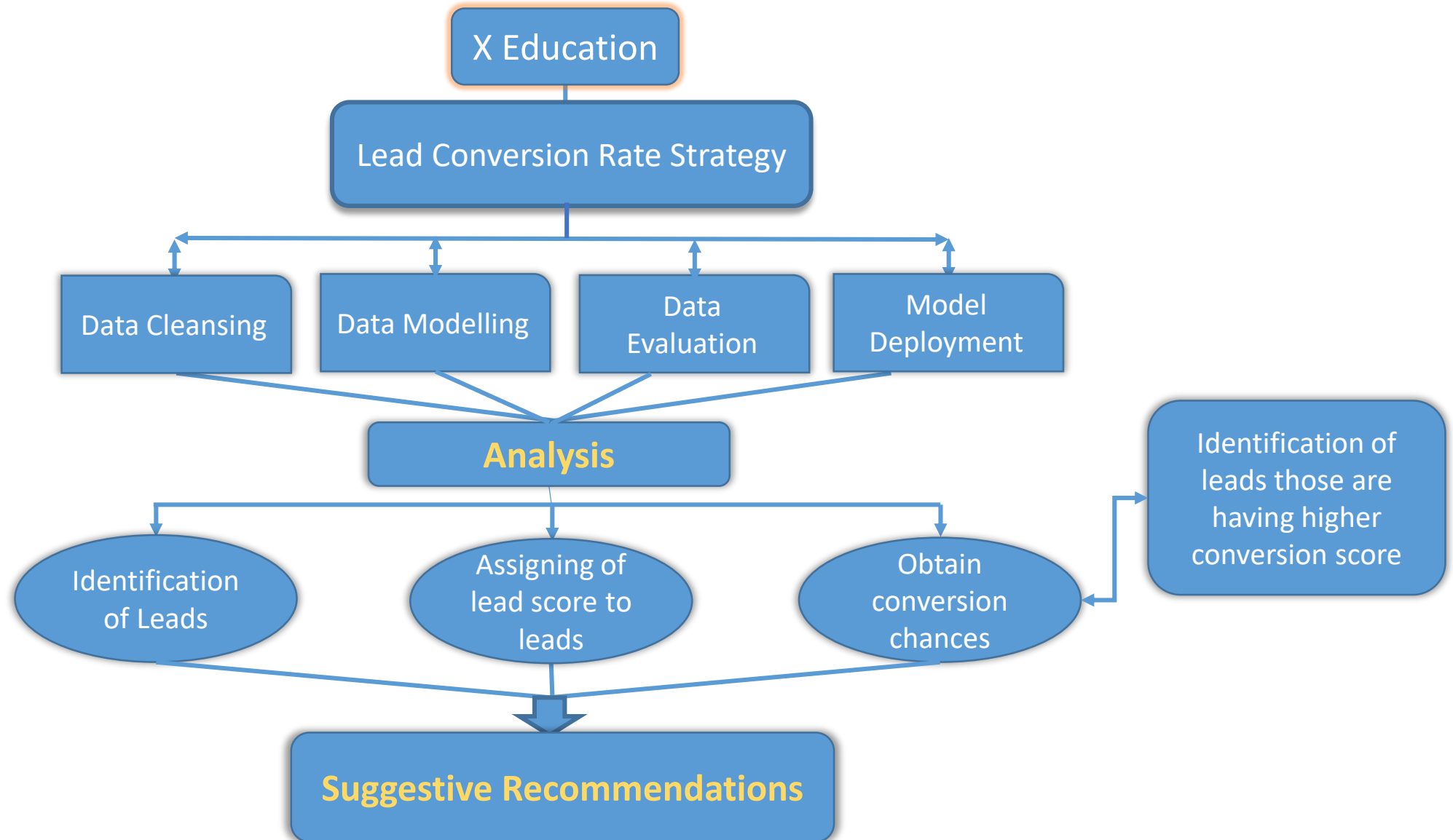
Group Name:

1. Gayathri Madhira
2. Apurva Deshpande
3. Karthik
4. Jayashri M

## SYNOPSIS

X Education, an Education Company sells online courses to industry professionals. Industry professionals who seek for pursuing any of online courses browse for courses in their website. The marketing strategy of X Education is made available of their courses in several website and search engine like Google. The people who have browsed for information of online courses offered by X Education fill up their form for respective online courses and these people are classified as Lead by X Education. Through the sales team, tapping of these Leads will lead into progressive conversion rate. Irrespective of this rigorous process, company's lead conversion rate is very poor. To make this more effective, company wants to identify the most potential leads such as Hot Leads. Their approach is to successfully identifying of these hot leads which potentially bring their conversion rate to boom and the sales team can tap and communicate with the potential leads rather than making calls to everyone.

# METHODOLOGY



# METHODOLOGY

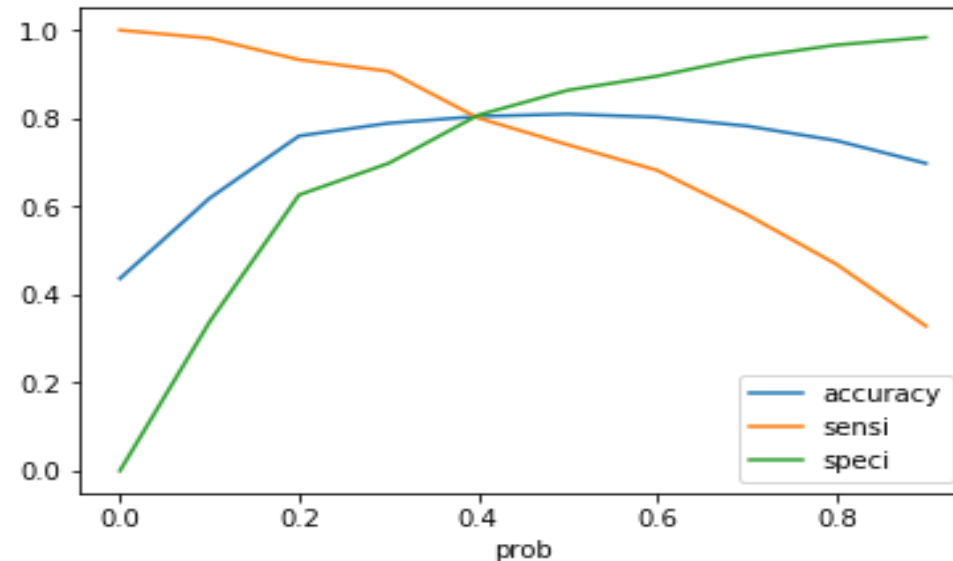
Data cleansing process was done to ensure data is correctness, consistent and useable by identifying any errors or corruptions in the data, corrected them to prevent the errors for building logistic regression model. Missing value treatment, outlier treatment and removing of erroneous values were also involved at initial phase. Data standardize process was followed to ensure a good point of entry and reduce the risk of duplication. Created dummy variables for categorial variables to increase information from the data and to reduce the number of variables originally present in the data set. At a later stage Data Modelling was followed by building a logistic regression model to assign a lead score between 0 and 100 to each of the leads that can be used by X Education to target the potential leads. A higher score which are Hot leads and lower score which lower scores were arrived at to obtain conversion scores to assist the X Education in tapping the Hot Leads which are mostly potentials converts.

# METHODOLOGY

According to the requirement of X Education we have built a model wherein we have assigned a lead score to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion rate. This would help the CEO of X Education to tap the right potential lead who are having high conversion rate. Test-train split of the data and standardisation of the scales of continuous variables was followed. After all of this was done, a logistic regression model was built in Python using the function GLM() under statsmodel library. This model contained all the necessary variables some of which had insignificant coefficients. Hence, some of these variables were removed first based on an automated approach – RFE and then a manual approach based on the VIFs and p-values. From the confusion matrix, correctly predicted variables were arrived at. For our model, we got an accuracy of about 80% which seemed good but if we look at the confsion matrix, there are lot of misclassifications.

# METHODOLOGY

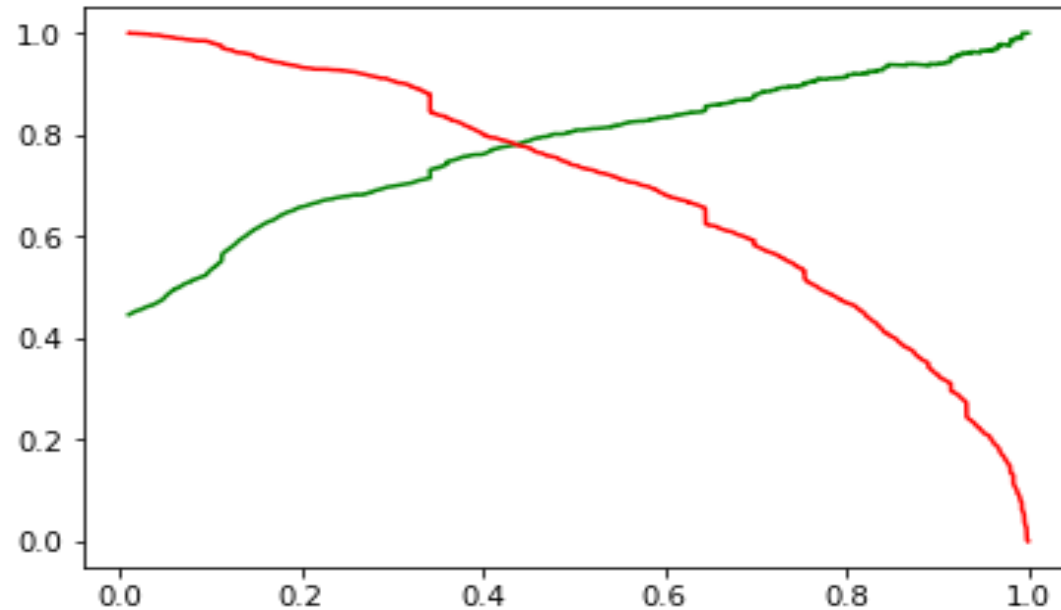
Hence, we calculated Sensitivity and Specificity and found that both Sensitivity and Specificity are about 80%. To obtain above the cut-off point was optimised in order to get a decent value of sensitivity and plotted RoC curve. Then we plotted the accuracy, sensitivity and specificity and got the following curve:



From this, we concluded that the optimal cut-off for the model is around 0.4 and obtained decent values of Accuracy (80.4%), Sensitivity (80%) and Specificity (80.7%) .

# METHODOLOGY

Similarly we plotted trade-off curve between precision and recall.



After worked out with the metrics, we chose a cut-off point of 0.4, we went ahead and made predictions on the test set and got decent values and decided to be have it as out final model.

# RECOMMENDATIONS AND CONCLUSION

As a concluding remarks our group would like to recommend the following:

- To retain the leads to take up more and more online courses of X Education, the company has to offer free bonus training, discount for another course, special offer for consulting, offering ultimate resume makeover course.
- Offer advanced management training like leadership, negotiation + more as a supplement to course as it attract industry professionals more as a relevant skill set for reaching out to the path breaking phase in his/her career.
- Dynamic offers to be offered to industry professionals according to the changing industry trends i.e. the booming trend is analytical development.