Summary

Missing out on a substantial proportion of Leads that are generated by X Education, is making the business model less efficient, as a result the company is looking to identify ways in which they could be more focussed. Essentially improve the conversions from a meagre 30% to a high of 80%.

X Education has a robust database of the leads generated in the past. This data can be used build a deeper understanding of their leads and also to identify ways in which they could target their Leads better and be more productive in lead conversions.

The Approach here is two pronged –

1. At the very first stage is to help identify leads as Hot Vs Cold as soon as they reach out to X Education. This will help prevent wastage of time and resource on leads that are less likely to convert and also target the prospective ones with a more rigorous process.
2. Secondly, as an additional output from the Model, X Education can better ways and means to reach the right Leads. This will help allocate marketing budgets better.

After going through the data file, it looked more like a case of Logistic regression, where we could build a model to assign a score to each lead. Based on this score, the Company would know which customers are the top ones to actively engage with. A score of greater than 0.46 was found to be a positive indicator towards Lead conversion.

A rigorous data cleaning and EDA was performed before model building.

It was also ensured that the difference between Sensitivity and Specificity was low. This would help on reducing False Negative and False positive errors. Although the Model accuracy hovers around 75%, the specificity and sensitivity scores are also within close range of this. This makes the model quite robust.

Also ,the Accuracy of the model does not drop from Train Data to Test Data, significantly. They operate in a similar range.

The model output can be used for the following-

1. Firstly, the Model will help X Education with a score against each Lead. This will be key to identifying a lead as Hot or Cold. Leads with >0.46 cutoff, could be considered as hot Leads.

2. Secondly, based on the key drivers, X Education can focus on very specific marketing activity on certain websites or / and also decide means of targeting these lead (like specifically only tele -calling and no charting allowed)

    • The top 3 variables in our model which contribute highly to the lead generation are

Total time spent on Website (3.4)

Last Notable Activity of a Phone Conversation (2.9)

Lead Source being Welingak Website (2.3)

With this, we can infer that more emphasis should be provided if there is more time spent on website and lead source as Welingak website. Also, it's evident that only interested and high chance of conversion students opt for phone calls.

More promotions and campaigning can be concentrated for such students