Summary : Lead Scoring Case Study

**Problem Statement :**

An education company named X Education get leads via various sources and methods. Their current conversion rate is approx. 30%, i.e. 30% leads get converted to real subscription for their program. Need to create a ML model which predicts and assign a Lead Score so that Sales team can target only “Hot Leads”, i.e. which have extremely high chances of getting converted. Target is to achieve 80% positive conversion ratio by optimising resources.

**Analysis Approach:**

* Understand data and explore various features of it.
* Data cleaning
* Handle missing values,
* Impute for unwanted values
* Handling outliers. Etc.
* Prepare the data for modelling
* Handling imbalance in data
* Converting categorical columns to Binary
* Creating Dummy variables for multi-layered categorical columns
* Split data to train-test
* Scaling the data
* Building the model
* Train the model
* Test the model

**Model outcome :**

We assigned a LEAD SCORE to each lead. A lead score tell the probability of lead getting converted in an actual subscription. Higher the lead score, more are chances of lead getting converted.

After analysis we found optimal probability cut-off as 0.35.

This means if lead score is more than 35, lead is likely to be converted.

Average lead score of converted leads => 69

Average lead score of not-converted leads => 15