Problem Statement, Understanding and Formulating ML solution

# Business Problem

An education company “***X Education***” in the business of selling online courses to industry professionals.

Company markets their courses on websites and search engines.

Once people land on the website and fill the form mentioning their personal information such as email address or phone number they become “***lead***”.

These “***leads***” are then contacted by the company’s sales representative through calls or e-mails. An average of **30%** of such leads are converted to ***paying customers***.

Since the conversion rate is low the company wants to save resources exhausted on the **70%** non-convertible client and hence wants to target only the clients that are very highly probable to become a client also termed as “***Hot Leads***” . This in turn ensures the conversion rate goes high and lead conversion process turns more efficient.

Company wants to help us improve the efficiency by assigning score that is directly proportional to the conversion chance of the leads. The scores are expected to be between ***0-100****.*The **CEO** has given the ballpark target of improving the lead conversion rate to **80%** from the current **30%**

**Business metric:** lead conversion rate

Problem Mapping

The task is to identify or classify hot leads. Along with, classification a score between 0-100 is to be assigned such that higher score indicates more likeness for lead conversion, which can be achieved with the help of a classification model that assigns a probability score to each customer.

**Major Data source:** data tracked by client for each lead whether converted or non-converted to customer which covers different data points with respect to customer.

Interpretability is an important parameter. Accurate classification is important. FN and FP should be least and TP and TN should be high

**Data science metric:** Sensitivity/Recall/True Positive Rate(TPR), True Negative Rate(TNR), AUC-ROC, Precision, Confusion Matrix