**Lead Scoring Case Study Summary**

**Problem Statement**: X Education sells online courses to industry professionals. X Education needs help in selecting the most promising leads, i.e. the leads that are most likely to convert into paying customers. The company needs a model wherein you a lead score is assigned 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 chance. The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%

**Solution Summary**: **Step1**: Reading and Understanding Data.

**Step2**: Data Cleaning: In thi step we try and segregate the available data according tour needs and make sure the required data is kept and the uncessary data is droped

**Step3**: Data Analysis In this step we use EDA to determine how the data is constituted and we also drop duplicate and null values.

**Step4**: Creating Dummy Variables Here we create dummy variables

**Step5**: Test Train Split: This step is used to divide the data into train and test

**Step6:** In the next step we eliminated unwanted data

**Step7**: We then tried plotting the ROC

**Step8**: Then we plotted the probability graph for the ‘Accuracy’, ‘Sensitivity’, and ‘Specificity’. The intersecting point for the cutoff point was found out to be 0.4 We could observe ‘sensitivity=77.8%’, ’specificity=78.8%’.

**Step10**: Computing the Precision and Recall metrics we got the cut off at 0.4