

## **JOB-A-THON MAY 2021 : CREDIT CARD LEAD PREDICTION**

### **Aim:**

The objective of this competition was to predict whether the customer is interested in buying Credit Card or not given their details and relationship with the bank.

### **Approach:**

Here, based on the input the prediction was to made on the *Is\_Lead* column which determines whether to give credit to a particular user or not. This was a classification problem.

### **Data Pre-processing & Model Selection:**

Firstly, I filled all the missing values using filling all of them by using '*ffill*' method. After fixing the null values, I convert all the categorical variables which in string to 0's and 1's using Label Encoders. Then using these values, I created a Decision Tree Classifier which gave a score of 62%.

After this I also included the Age and Gender as an input and re-trained the model which gave me the accuracy of 66%.

After this I also made some pre-processing on Region\_Code and also included it into the training. On, including all of them I build a Random Forest Classifier with few changes in default parameters which gave me a score of 70%.

After doing some research I found boost classifier among which I tried XGBoost, LGBM boost and Cat Boost. From this LGBM boost gave the accuracy of 73%

### **Things I tried but didn't work:**

Splitting the 'yes' and 'no' values from *Is\_Lead* so as to improve the training.

Pre-processing to splitting the region code to make like a pivot table to improve accuracy.