Credit Card Lead Prediction

**Approach:**

First, it’s necessary to accurately define the data problem that is to be solved. Here, our main task is to identify customers that could show higher intent towards a recommended credit card.

I have combined the train and test data to get more insights from the data and also to understand if we have any missing data or not.

**Data-preprocessing / feature engineering:**

We can find that we have a large number of missing values in **Credit\_Product** which needs to be treated. So, I have imputed the column with values which is less frequent which is ‘**Yes’.**

I have dropped the columns which are irrelevant, i.e, ‘**ID**’ which is a unique value and which won’t help us in training the model.

**Final model:**

I have used the KFold and StratifiedKFold libraries to train and evaluate the data with folds=9. I trained LGBMClassifier, XGBClassifier and CatBoostClassifier models with the given training data and evaluated that on the given test data. The final **Is\_Lead** values are an average of the values generated from the above 3 models.