**Credit card Classification Project**

Credit card is now the day most popular; every bank releases their credit card for customers to spend with a limit and then pay in installments. In this project, I have been doing an analysis of a Credit card default payment dataset and I have built machine learning models to predict the target variable. I have been using machine learning algorithms like Logistic, Decision Tree, Random Forest, KNN, and XG Boost.

**Steps:**

First, I have to know about the dataset like

* Import different important libraries
* Load dataset
* There are 30000 rows and 25 columns
* There is no duplicate or missing data in the dataset.

Then understand variables or features

* Check the column’s name and type
* Check the statistical description of the continuous features.

I did some exploratory data analysis of the dataset to find some patterns or trends.

* I found that 22% of customers have default payments next month.
* Pay amount is the most influential factor for default payment next month.
* Young people of age less than 35 uses credit card most.

Then I did some feature engineering.

* The Target variable was imbalanced, so I used SMOTE to balance the target variable.
* Converted categorical features into continuous features.

Now my dataset is ready for model implementation.

* I have implemented 5 different machine learning models to predict the target variable.
* I first trained my model and then tested it on the test dataset.
* For the performance, I have used accuracy, precision, recall, and the f1-score of each model.

**Result:**

All the performance values of each model are given on the test dataset.

* The Logistic classifier has an accuracy of 61.2%
* The Decision Tree has an accuracy of 74.52%
* The Random Forest has an accuracy of 78.68%.
* KNN classifier model has an accuracy of 80%
* XG Boost has an accuracy of 82.71% which is the highest of all models.

I can clearly say that XG Boost is the best and all performance metrics are highest among all models.