**Project Name:** Loan Repayment Prediction

**Github Link:** https://github.com/utkrisht2000/Loan-Repayment-Prediction.git

**Why was this project created?**

Many people struggle to obtain loans from reputable sources like banks due to weak credit records. These persons are typically college students or unemployed adults who may lack the expertise necessary to support the legitimacy of the unnamed lenders. These borrowers may be victimized by unreliable lenders that charge exorbitant interest rates or include secret clauses in contracts. There are several different methods to gauge or forecast the borrower's capacity to repay the loan besides using their credit score. Since a working adult has more steady income and cash flow, employment can have a significant impact on the person's capacity to repay the loan. The examination of the ability to repay a loan might also benefit from other variables like real estate and marital status.

**What problem is it solving?**

To reduce the risk of loan payment, banks must assess and forecast the loaners' capacity to repay their loans. Based on the borrower's situation, including job and credit history, this system processes the loan request. However, certain borrowers' capacity to repay loans, such as students or those without credit records, may not be adequately assessed under the current system. We develop a model based on the relevance score of the characteristics, evaluate the data, and pick the most recognisable aspects to forecast the ability of the loaner to repay the loan in order to accurately estimate the repayment capacity of all groups of individuals.

**Entire explanation of project**

* **PROPOSED APPROACH**

The project's objective is to forecast the borrowers' capacity to repay the debt using variables other than their credit histories. A classification problem with two classes can be used to describe it. We'll outline the techniques we utilised to pre-process the data and the machine learning techniques we'll employ to address the issue. We factorise these characteristics using label encoding, which translates the string values to category values, each represented by an integer, to prevent classification biases towards certain features. We test several machine learning algorithms for the purpose of forecasting borrower repayment capacity. Among the machine learning algorithms are decision trees, random forests, and logistic regression. Some of the methods were taught to us in class, while others were discovered online and may perform well on the dataset.

We divided the target (y) and feature (X) dataframes into training (80%) and testing (20%) sets. The training set is used to create the classification model, and the testing set is used to assess the model's effectiveness. The objective of the project's unsupervised learning component is to get some important insights into the data's structure and maybe classify the various kinds of loan applicants in our dataset. In order to create alternative prediction models for various categories of borrowers, we sought to determine whether different groups of borrowers had distinguishing traits. We experimented with Random Forest and Decision Tree as machine learning techniques.

Algorithm for creating next word prediction model :

**Step 1:** Dataset is imported

**Step 2:** The data is preprocessed, and the dataset is split into training and testing.

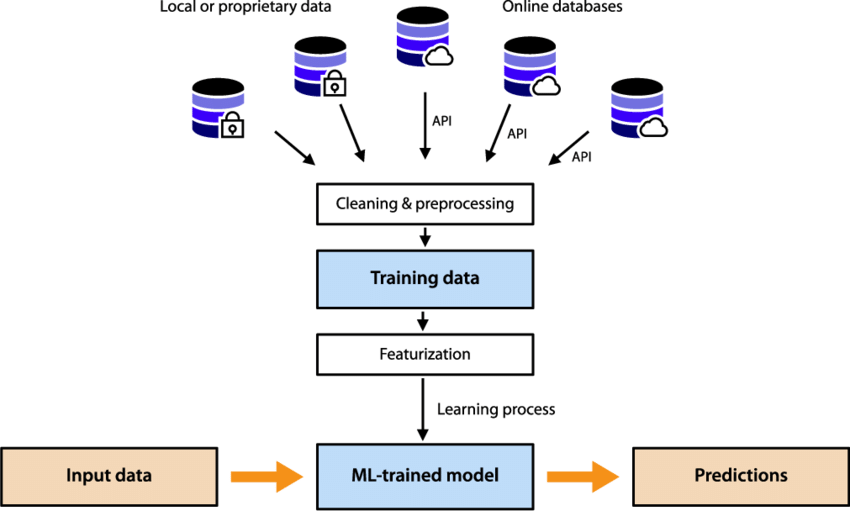
**Step 3:** Feature Extraction

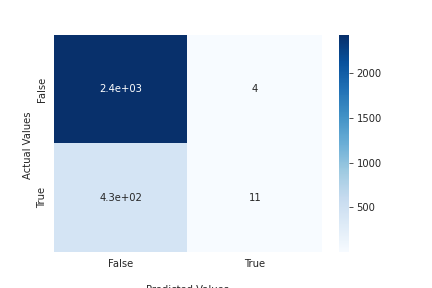
**Step 4:** Use Decision Tree

**Step 5:** Create Random Forest Model and Train the model

**Step 6:** Calculate Accuracy

* **DATA FLOW DIAGRAM**



* **RESULT**
* **CONCLUSION**

On a very difficult dataset, we showed how machine learning algorithms may be used to forecast a person's capacity to repay a loan. We demonstrated that data pre-processing, a careful selection of dataset balancing strategies, and machine learning algorithms are all crucial for achieving the optimum results. On our dataset, Random Forest performs admirably, and Decision Trees are very useful. Our accuracy now stands at 84.00%.