**Smart Lender - Applicant Credibility Prediction For Loan Approval**

**IDEATION:**

### ****XGBoost**** ****for Loan Prediction using Machine Learning****

[Boosting](https://www.projectpro.io/article/bagging-vs-boosting-in-machine-learning/579) is a method that combines individual models in an ensemble manner to gain higher performance gain. AdaBoost and Stochastic Gradient Boosting are the most popular boosting algorithms where higher weights are assigned to wrong classified instances during training. At the same time, SGB adds randomness as an integral part of training. Extreme Gradient Boost (XGBoost) is an improvement over Gradient Boost and is very popular in binary [classification algorithms](https://www.projectpro.io/article/7-types-of-classification-algorithms-in-machine-learning/435). The decision trees are built in parallel in XGBoost than in series, giving it an edge over normal Decision Trees and Boosting algorithms.

### ****Support Vector Machine for Loan Prediction using Machine Learning****

Support Vector Machine (SVM) is a supervised machine learning algorithm that generates a hyperplane (a decision boundary) to separate classes even in a high-dimensional vector space. It can capture different non-linear relationships between the features and the target variable. It decides a class for a sample based on the sign of w[T]+b. In the equation, w (weights) represents the negative and positive hyperplane margin, and b is the bias. SVM is particularly useful in loan prediction because this task usually has several features that need to be considered for the final decision

### ****Random Forest for Loan Prediction using Machine Learning****

The random forest algorithm improves the flexibility and decision-making capacity of individual trees. It is another machine learning algorithm incorporating the ensemble learning theorem as its foundation, combining results from various decision trees to optimize training. In some use cases of loan and credit risk prediction, some features are more important than the rest or, more specifically, some features whose removal would improve the overall performance. Since we know the fundamentals of decision trees and how they choose features based on information gain, random forests would incorporate these benefits to give superior performance.

### ****Loan Credibility Prediction System Based on Decision Tree Algorithm****

The proposed model focuses on predicting the credibility of customers for loan repayment by analyzing their behavior. The input to the model is the customer behavior collected. Based on the output from the classifier, decision on whether to approve or reject the customer request can be made. Decision Tree Induction data mining technique is used to generate the relevant attributes and also make the decision in the model.

### ****Loan Credibility Prediction System Based on KNN Algorithm****

KNN is a simple algorithm, based on the local minimum of the target function which is used to learn an unknown function of desired precision and accuracy. The algorithm also finds the neighborhood of an unknown input, its range or distance from it, and other parameters. It’s based on the principle of “information gain”—the algorithm finds out which is most suitable to predict an unknown value.