PROPOSED SOLUTION

Problem statement:

To design and implement the system using machine learning and data mining to predict the probability of the user to get loan or not from bank to improve the accuracy and to minimize the frauds. Banks, Housing Finance Companies and some NBFC deal in various types of loans like housing loan, personal loan, business loan etc in all over the part of countries. These companies have existence in Rural, Semi Urban and Urban areas. After applying loan by customer these companies validates the eligibility of customers to get the loan or not. This project provides a solution to automate this process by employing machine learning algorithm. So the customer will fill an online loan application form. This form consist details like Gender, Marital Status, Qualification, Details of Dependents, Annual Income, Amount of Loan, Credit History of Applicant and others. To automate this process by using machine learning algorithm, First the algorithm will identify those segments of the customers who are eligible to get loan amounts so bank can focus on these customers.

Idea:

A Machine Learning algorithm is to be used in order to construct a robust and efficient software algorithm that classifies individuals based on different characteristics (Gender, Education, Number of Dependents, Marital Status, Employment, Credit Score, Loan Amount, and others) whether they would be eligible for a loan or not.

Ensemble learning using multiple ML models is the solution that is going to be used. Ensemble modeling is the method of running two or more associated but different models and then combines the results into a single score to improve the accuracy of predictive data and data mining applications. In machine learning, ensemble methods use several algorithms to get better predictive performance. The different ML models that can be used are KNN, Decision tree, Random forest, Xgboost.

Novelty:

Ensemble modeling combine multiple facts to form a better result. This method of prediction has been shown to enhance forecasts when compared to a single model- based approach. The main benefits of Ensemble models are: Better Forecasting ,More Constant model , Better results , Reduces error. All these factors make the project unique.

Social Impact/Customer Satisfaction:

Now a day's bank plays a vital role in market economy. The success or failure of organization largely depends on the industry's ability to evaluate credit risk. Banks have many products to sell in our banking system, but their main source of income is their credit

lines. As a result, they are likely to profit from the interest on the loans they make. Loans, or whether customers repay or default on their loans, affect a bank's profit or loss. The bank can minimize its Non-Performing Assets by forecasting loan defaulters. Because precise predictions are crucial for maximising earnings, it's essential to look at the different methodologies and compare them.

Business Model(Financial Benefit):

The bank can minimize its Non-Performing Assets by forecasting loan defaulter.Furthermore, automation of the loan approval processes opens new financing opportunities for small businesses and individuals. These previously suffered from limited access to credit, due to the high cost of human involvement in the process. Ultimately, automation of this process carries the potential to reduce human bias and corruption, making access to credit fairer for all.