

# Smart Lender - Applicant Credibility Prediction for Loan Approval

## Literature survey

- [1] Sai Aparna Vangaveeti et al, In this paper, Logistic Regression algorithm had been used for predicting whether the loan is approved or not. They have used various input variables to get the efficient output. The output will be in form of binary (0 or 1). If the output is 0, then there is no approval. If the output is 1 means it is approved
- [2] Kumar Arun et al, This paper's major goal is to forecast whether or not issuing a loan to a specific person will be secure. There are four components in this paper: (i)Data Collection (ii) Comparison of machine learning models on collected data (iii) Training of system on most promising model (iv) Testing. They have used six machine learning methods decision tree, random forest, support vector machine, neural network, adaboost, linear model
- [3] Yamuna B et al, In this paper, they have used five different machine learning model to find the best fitting model. The five different models are linear regression, random forest, support vector machine, decision tree classifier, extreme gradient boost (XG Boost). Among all these machine learning algorithms XGBoost proved to be the best with accuracy 0.82
- [4] Rekha A.G et al, In this paper, to analyze the relationship between financial transactions and loan defaults, two methods based on LSTM (Long Short Term Memory) and a hybrid neural network architecture are proposed. 94% accuracy is attained via bidirectional LSTMs with hybrid architecture. This work's utilization of hybrid neural network architectures offers the right path toward developing an early warning system through online loan default prediction
- [5] Kumar T et al, In this paper, data from former clients of several banks who had loans accepted based on a set of criteria were used. The machine learning model is trained on the record to get reliable results. To forecast loan safety, the logistic regression approach is applied. The data is initially cleaned in order to remove any missing values from the data set. A data set of 1500 examples with 10 numerical and 8 categorical features were used to train the model. When crediting a loan to a customer, a number of factors,

including the customer's CIBIL Score (Credit History), business value, and customer assets, have been considered

- [6] Vaidya, This paper, discusses logistic regression and its mathematical representation. In order to implement the predictive and probabilistic approaches to a specific problem of loan approval prediction, his study uses logistic regression as a machine learning technique. This is to decide whether to approve a loan for a set of records belonging to an applicant. It also discusses a few other practical applications for the Machine Learning mode
- [7] Chandan Soni et al, In this paper, a loan prediction system is presented that aids organizations in choosing whether to approve or deny loan requests from clients. The input variables such as loan ID, Marriage, loan amount, gender etc., have been sent to train the model. The prediction is made using the Decision Tree Algorithm
- [8] Tanvir Anzum et al, In this paper, in order to predict fraudulent loan requests from clients, they have used six machine learning algorithms such as Decision tree, Support vector machine, Random forest, K closest neighbors, Ada-Boost, and Logistic regression. The K-Nearest Neighbors algorithm provided accuracy of 83.75%, which was superior to the other five machine learning techniques
- [9] Jayapandian N et al, In this paper, they have used two different kind of datasets. One is used for training the model and the other dataset is used for testing the model. They used random forest, machine learning algorithm to train the model. The suggested random forest model offers a higher level of accuracy. This approach offers a level of accuracy that is 28% higher than that of conventional prediction
- [10] Suresh Kumar et al, The primary goal of this paper is to decide whether or not the loan offered to a specific individual or organization should be approved. They have used two machine learning algorithms one is random forest and another one is decision tree classifier. The parameter setting for decision tree are Min Split=20, Max Depth=30, Min Bucket=7. The parameter setting for random forest are Number of trees=450, number of variables=8