Loan Approval Prediction using Machine Learning

Krishna Patel[1], Manthan Patel[2], Prof. Ajaykumar T. Shah[3]
Dept of Computer Engineering
[1,2]Student, Alpha College of Engineering and Technology
[3] HOD, Alpha College of Engineering and Technology

Abstract- *With the enhancement in the banking sector lots of people are applying for bank loans but the bank has its limited assets which it has to grant to limited people only, so finding out to whom the loan can be granted which will be a safer option for the bank is a typical process. So in this paper we try to reduce this risk factor behind selecting the safe person so as to save lots of bank efforts and assets. This is done by mining the Big Data of the previous records of the people to whom the loan was granted before and on the basis of these records/experiences the machine was trained using the machine learning model which give the most accurate result. The main objective of this paper is to predict whether assigning the loan to particular person will be safe or not. This paper is divided into four sections (i)Data Collection (ii) Comparison of machine learning models on collected data (iii) Training of system on most promising model (iv) Testing*

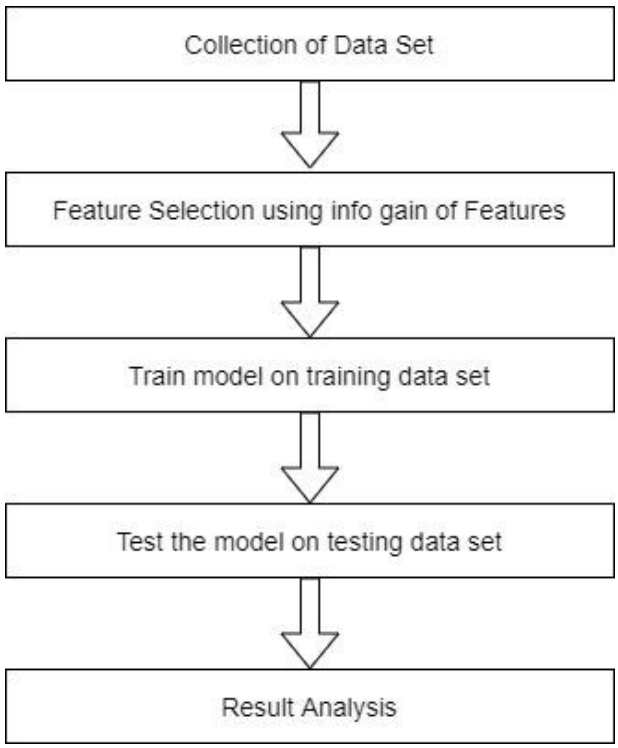
Keywords- Training the models, testing, prediction, Loan, Machine Learning

I. INTRODUCTION

Distribution of the loans is the core business part of almost every banks. The main portion the bank's assets is directly came from the profit earned from the loans distributed by the banks. The prime objective in banking environment is to invest their assets in safe hands where it is. Today many banks/financial companies approves loan after a regress process of verification and validation but still there is no surety whether the chosen applicant is the deserving right applicant out of all applicants. Through this system we can predict whether that particular applicant is safe or not and the whole process of validation of features is automated by machine learning technique. The disadvantage of this model is that it emphasize different weights to each factor but in real life sometime loan can be approved on the basis of single strong factor only, which is not possible through this system. Python is platform used to configure the data.

II. DATA SET

2.1 Loan Prediction Methodology



2.2 Machine Learning Methods:

Six machine learning classification models have been used for prediction of android applications .The models are available in R open source software. R is licensed under GNU GPL. The brief details of each model is described below.

2.2.1 Decision Trees (C5.0):

The basic algorithm of decision tree [7] requires all attributes or features should be discredited. Feature selection is based on greatest information gain of features. The knowledge depicted in decision tree can represented in the form of IF-THEN rules. This model is an extension of C4.5 classification algorithms described by Quinlan.

2.2.2 Random Forest (RF):

Random forests [8] are a group learning system for characterization (and relapse) that work by building a large number of Decision trees at preparing time and yielding the class that is the mode of the classes yield by individual trees. Every user should be comfortable for the working of the

known as a basic computer and net browser. They must have basic knowledge of English Language. User has to login one time. User can select the desired person by selecting categories. User must have some knowledge of how to use any websites. They have been some creating account of basic needs.

2.2.3 Linear Models (LM):

The Linear Model [10] is numerically indistinguishable to a various regression analysis yet burdens its suitability for both different qualitative and numerous quantitative variables.

III. FUTURE ENHANCEMENT

Many banks/financial companies approves loan after a regress process of verification and validation but still there is no surety whether the chosen applicant is the deserving right applicant out of all applicants. Through this system we can predict whether that particular applicant is safe or not and the whole process of validation of features is automated by machine learning technique.

IV. CONCLUSION

From a proper analysis of positive points and constraints on the component, it can be safely concluded that the product is a highly efficient component. This application is working properly and meeting to all Banker requirements. This component can be easily plugged in many other systems. There have been numbers cases of computer glitches, errors in content and most important weight of features is fixed in automated prediction system, So in the near future the so – called software could be made more secure, reliable and dynamic weight adjustment .In near future this module of prediction can be integrate with the module of automated processing system. the system is trained on old training dataset in future software can be made such that new testing date should also take part in training data after some fix time.

V. ACKNOWEDGMENT

We express our sincere thanks to Prof. Ajaykumar T. Shah Head of Department of Computer Engineering, Alpha College of Engineering and Technology for their Support and guidance for this project and care taken by them in helping us to complete the project work successfully.

REFERENCES

[1] Rattle data mining tool: available from https://rattle.togaware.com/rattle-download.html
[2] Ekta Gandotra, Divya Bansal, Sanjeev Sofat 2014, ‘Malware Analysis and Classification: A Survey’available from http:// www.scirp.org/journal/jis.
[3] For python Programming:
[4] For Designing machine learning , python and Others: https://www.w3schools.com/
[5] Programming and Designing: www.simplilearn.com/PGP/AIML
[6] Introducing to machine learning with python: A Guide for Data Scientist Originally published: October 2017