Predicting Personal Loan Approval Using Machine Learning

1. INTRODUCTION

Loan Distribution is the main business part of many banks. The main portion of banks income comes from the loan distributed to customers. These banks apply interest on loan which are distributed to customers.The main objective of banks is to invest their assets in safe customers. Up to now many banks are processing loans after regress process of verification and validation. But till now no bank can give surety that the customer who is chosen for loan application is safe or not. So to avoid this situation we introduced a system for the approval of bank loans known as Loan Prediction System Using Python.

Loan Prediction System is a software which checksthe eligibility of a particular customer who is capable of paying loan or not. This system checks various parameters such as customer’s martial status, income, expenditure and various factors. This process is applied for many

customers of trained data set. By considering these factors a required model is built. This model is applied on the test data set for getting required output. The output generated will be in the form of yes or no. Yes indicates that a particular customer is capable of paying loan and no

Prediction of Loan Approval using Machine Learning

Approach

Machine learning is a phenomenon in which analytical model is build from the trained model.

This model is applied on test data for providing of the

accurate results.

Here the author used three algorithms for prediction of

loan. They are

1. K Nearest Neighbor

2. Decision Tree

3. Random Forests

The main purpose of this report is to provide

immediate and accurate results for the approval of loan to

the eligible customers. In banking sector there will be n

number of people who apply loans. It is difficult to check

customer’s eligibility through paper work. The system can

provide accurate results for the n number of people.

Existing System

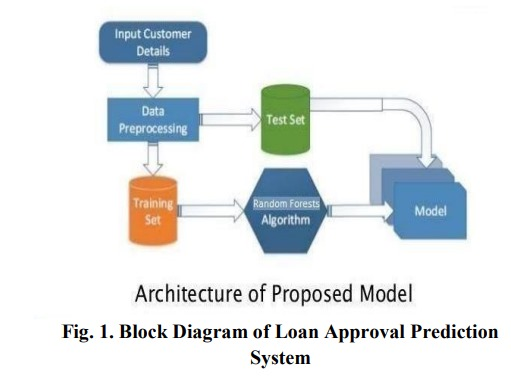
Till now loans are processed by various banks through pen and paperwork. When the large no of customers’ apply for bank loan these bank take lot of time to approve their loan.

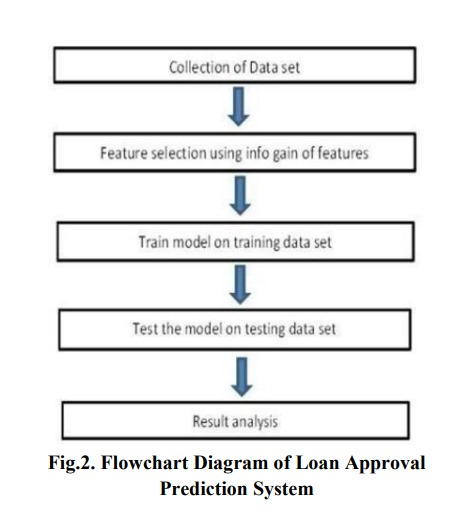
After approval of loan by the banks, there is no surety that the chosen applicant is capable of paying loan or not. Many banks use their own software’s for the loan approval. In existing system we use datamining algorithms for the loan approval; this is the old technique for the approval of loan. Mutiple data sets are combined and form a Generalised datasets, and different machine learning algorithms are applied to generate results. But these techniques are not up to the mark. Due to this huge banks are suffering from financial crises. To

resolve this issue we introduce a new way for approval of

loans

IMPLEMENTATION

We have used ensemble learning for building of the 

system

Working of Random Forest Algorithm

Random Forest Algorithm is follows rules of Decision Tree. The difference in them is decision tree algorithm gives the output by considering only one factor where as Random Forest Algorithm compares many no of decision trees and gives the result satisfying majority no of decision trees.

Journal of Innovation in Information TechnologyJan–Jun 2020@ ISSN: 2581-723X 23Random Forest Algorithm builds a strong model which satisfies the models many no of decision tree, this model is applied on the testing dataset for getting the required output. Firstly for building a model we require a training dataset, from that training dataset we consider the subset of training dataset known as bootstrap dataset1. This bootstrap dataset1 consists of set of variables. From these variables we consider only two variables from these two variables we make a root node for one variable which produces accurate output than the other variable. In this format we build a decision tree from the bootstrap dataset1. We consider another subset of dataset from the training dataset, let us say bootstrap dataset2. As we build a decision tree for the bootstrap dataset 1one in the same

way we build a dataset for the bootstrap dataset2. We have to follow these steps until we get many no of decision tree. After getting many no of decision trees we compare the entire decision tree and build a model which satisfies these entire decision trees. The obtained model is known as Strong Model. In this way a model is built from the training dataset.

This model is applied on the testing data set and it produces a required output. The obtained output is accurate because a strong model is build from many no. of decision treesIn this format, when we upload a training dataset then the system builds a model using Random forests Algorithm then when upload the testing dataset by using the model system provides the required output. The output consists of

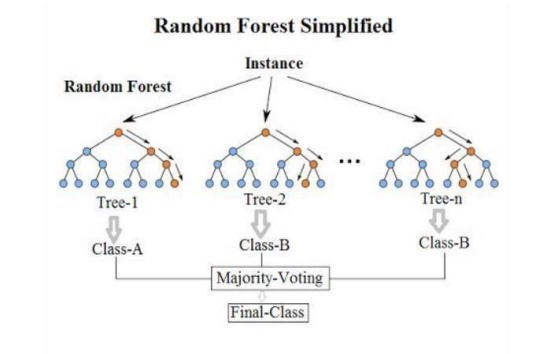
two class labels yes/no.

Yes indicates that the client is

eligible for approval of loan and no indicates that the client

is not eligible for the approval of loan. In this fashion the

system provides the required output.

 OUTPUT

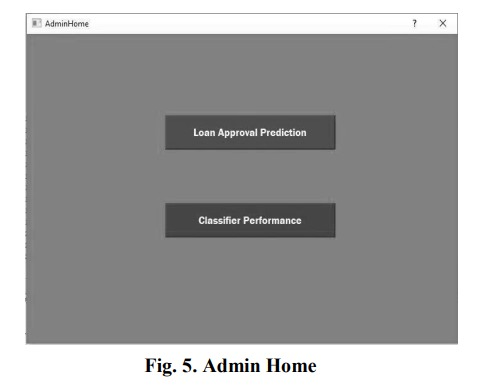
To protect system from unauthorized access, we have

created admin login module for security purpose. It consists

of username and password. By providing the valid username

and password we can access thesystem. The admin home consists of two options:

 Loan Approval Prediction

 Classifier performance Loan Approval Prediction is chosen for prediction

of loan status whereas classifier performance gives the

prediction results of three algorithms:

 K Nearest Neighbour

 Decision Tree

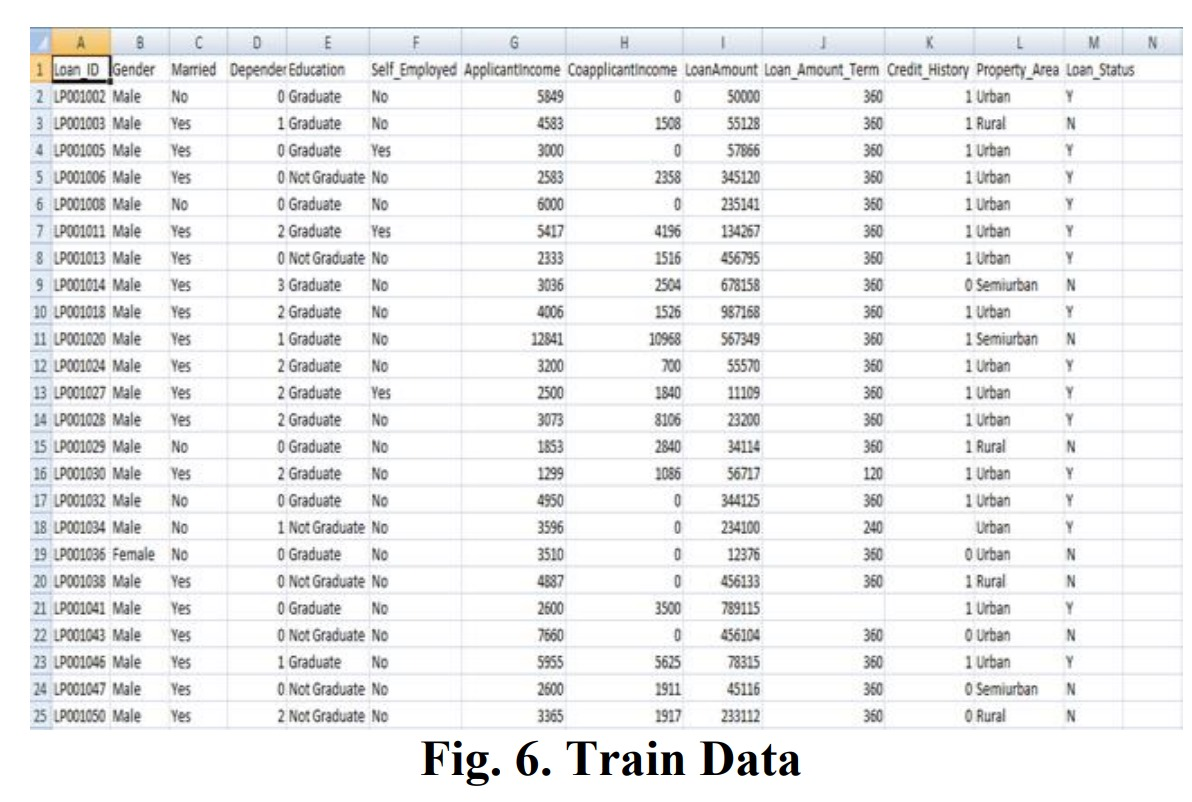
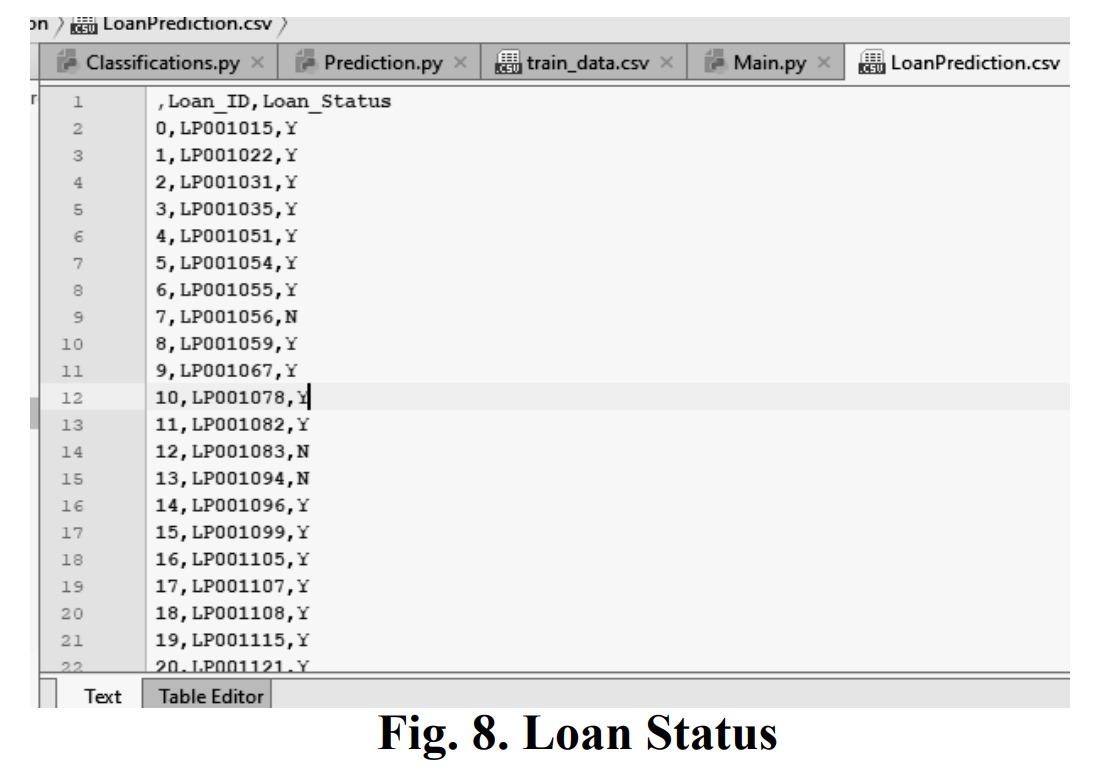
 Random Forest

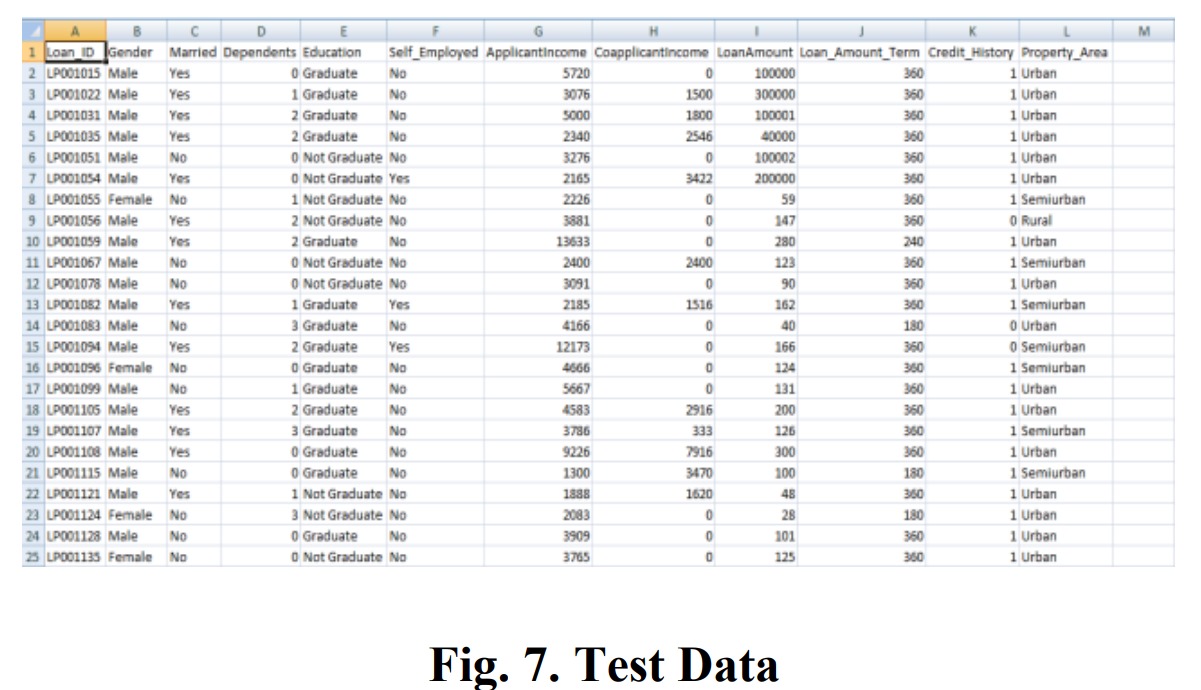
he above Fig. 7 represents the test data. The test

data consist of various attributes such as salary, martial

status, loan accounts except the loan approval status. The

loan approval status is obtained. When we deploy the test

data to the model which is build from the trained data.



The above Fig. 8 represents the loan status. The loan

status is obtained after the deployment of test data to the

model which is build from the trained data using Random

Forest Algorithm. The loan status consists of Customer id

and loan status. It indicates for a particular customer loan is

approved or not. If loan status is Y (Yes) then the customer

is eligible for approval of loan and if it is N (No) then the

customer is not eligible for approval of loan.

CONCLUSION

From the proper view of analysis this system can be

used perfect for detection of clients who are eligible for

approval of loan. The software is working perfect and can

be used for all banking requirements. This system can be

easily uploaded in any operating system. Since the

technology is moving towards online, this system has more

scope for the upcoming days. This system is more secure

and reliable. Since we have used Random Forest

Algorithm the system returns very accurate results. There

is no issue if there are many no of customers applying for

loan. This system accepts data for N no. of customers. In

future we can add more algorithms to this system for

getting more accurate results

indicates that the particular customer is not capable of

paying loan. Based on these factors we can approve loans

for customers.