

Loan Strategy Analysis System for Small Bank, Funding Clubs Industry and NGOs.

Name: Biprajit Namasudra

From: Assam University, Silchar

Abstract

In this report I have proposed a model for Loan strategies analysis system for small clubs and banks to help people open a startup for small business or help them during natural calamities or help them during food crisis. Here we provide loans for the people along with making sure that the organisation or any NGO provides help to the people along with not going through a loss. In this project we analyse the financial strength of the person or individual, their income pattern, what type of loan it is, for instance whether it's educational loan, small business, debt consolidation etc. and likewise we keep note on a numbers of attribute and apply suitable machine learning algorithm to predict the right person to whom we should provide the loan.

Problem Statement

The small banking and funding clubs undergoes several problems, which makes it difficult to decide to which particular sector of the population should we provide the loan amount especially while dealing with large population. Some of them are listed below:

1. Faster approval:

This is one of the most important and flaws of the traditional bank. The traditional banks and clubs take weeks to go through all the details of person and finally reach a conclusion to provide them a loan.

2. Employees:

It is another important factor to consider and often faced by the small banks. Number of employees and proper skilled and knowledge employees not just to offer a good services but also to properly advertise their scheme and marketing strategies. Unnecessary and unskilled employees is just a wastage of time and paper.

3. Large debt:

Small banks often faced large debt problems. Due to lack of proper knowledge of the schemes for which they are funding the loan, often they end up funding the organisation without proper financial background and for which they suffer from financial crisis and leads to uneven growth of their business.

4. Work Load:

While dealing with a large group of population the employees of the sectors has to undergo tremendous work load. This makes the job life of them frustrated one which lead to their bad attitude towards customer. Implementation of such a system can reduce their workload and will encourage them to offer a good services to more and more customer.

Market/Customer Need Assessment:

For a proper growth of the Indian economical structure the growth of the small funding structure and NGOs are to be considered an important one as they act as the backbone of the small educational institutions and business sectors. After the pandemic, the small sector is facing a huge loss regarding their business and finance.

On the view of customer needs, it secures the future for the bright students with poor financial background. Especially during the time of natural calamities such as the flood. In such cases the people lost contact with the outer world and such banks and small funding organisations are their only support. Helping such sectors with such ideal projects will be greatly beneficial for the people facing such disastrous situations.

Target Specification

The project idea is based on development of the NGOs and small banks who having do such works in helping the people and serving them during their time of needs.

The system will help them keep the customer satisfied, provide them with best service ever.

External Search

For analysis of system the sources I have used are :

- <https://www.projectpro.io/article/loan-prediction-using-machine-learning-project-source-code/632>
- <https://towardsdatascience.com/predict-loan-eligibility-using-machine-learning-models-7a14ef904057>

And the Data sources are collected from [LendingClub.com](https://lendingclub.com).

Import the libraries

```
In [2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
In [3]: df = pd.read_csv('loan_data.csv')
df.head()
```

```
Out[3]:
```

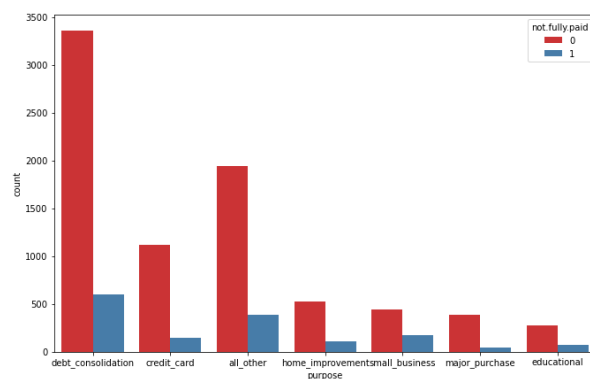
	credit.policy	purpose	int.rate	installment	log.annual.inc	dti	fico	days.with.cr.line	revol.bal	revol.util	inq.last.6mths	delinq.2yrs
0	1	debt_consolidation	0.1189	829.10	11.350407	19.48	737	5639.958333	28854	52.1	0	0
1	1	credit_card	0.1071	228.22	11.082143	14.29	707	2760.000000	33623	76.7	0	0
2	1	debt_consolidation	0.1357	366.86	10.373491	11.63	682	4710.000000	3511	25.6	1	0
3	1	debt_consolidation	0.1008	162.34	11.350407	8.10	712	2699.958333	33667	73.2	1	0
4	1	credit_card	0.1426	102.92	11.299732	14.97	667	4066.000000	4740	39.5	0	1

Benchmarking

The proposed model keeps track of several factors to produce a proper analysed outcome of whether we should or should not provide loan to the particular person. So, it keeps things simple and efficient at the same time. It is easy to use and also saves and workload and provides a better experience at the same time. It outputs a proper analysed graph showing the several important characteristics associated with a particular customer. Here is a figure for such analysis.

```
In [10]: plt.figure(figsize = (11, 7))
sns.countplot(x = 'purpose', hue = 'not.fully.paid', data = df , palette = 'Set1')
```

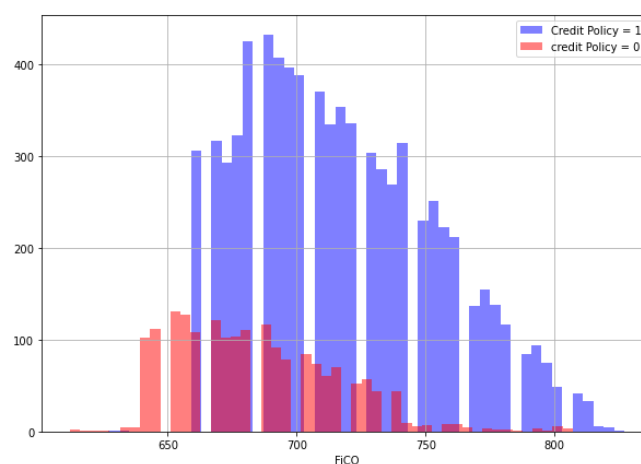
```
Out[10]: <AxesSubplot:xlabel='purpose', ylabel='count'>
```

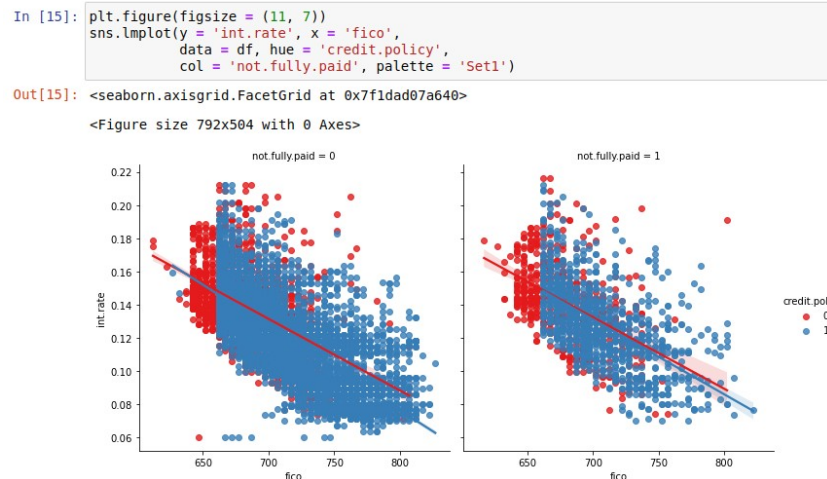


In [6

```
df[df['credit.policy']==0]['fico'].hist(alpha = 0.5, color = 'red',
bins = 50, label = 'credit Policy = 0')
plt.legend()
plt.xlabel('FICO')
```

```
Out[6]: Text(0.5, 0, 'FICO')
```





Here the above showing the different Exploratory Data Analysis Reports done on the various attributes of the customer, briefing what portion of the population should be considered.

Applicable Patents

Although there are many patents for development of the project, most useful one which I found are given below:

- US8458074B2
- US8775300B2
- US20150026035A1

External links for the patents:

<https://worldwide.espacenet.com/publicationDetails/biblio?CC=US&NR=2015026035A1&KC=A1&FT=D>

Applicable Regulations

- Convincing the owner to implement the system.
- Lack of technical knowledge for the user.
- Continuous data collections from the user.
- Taking care of customer interest, business plans and resources.

Concept Generation:

The recent flood in Silchar was such a disastrous. People were stuck in their home. Most of the people lost their stock and were suffering from financial crisis. At those tragic moments, several local groups and NGOs raised their hand to provide people with support. Using those funding they provided the people with food and most essential need is the water. Their efforts motivated me to develop such a model and which would be helpful regarding funding loans to the general public. Also such model requires less man power and simple installation and also much faster as compared to the local banks.

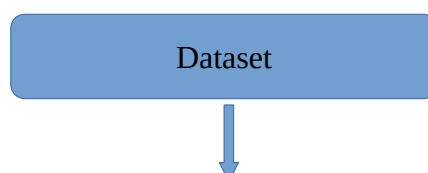
Concept Development:

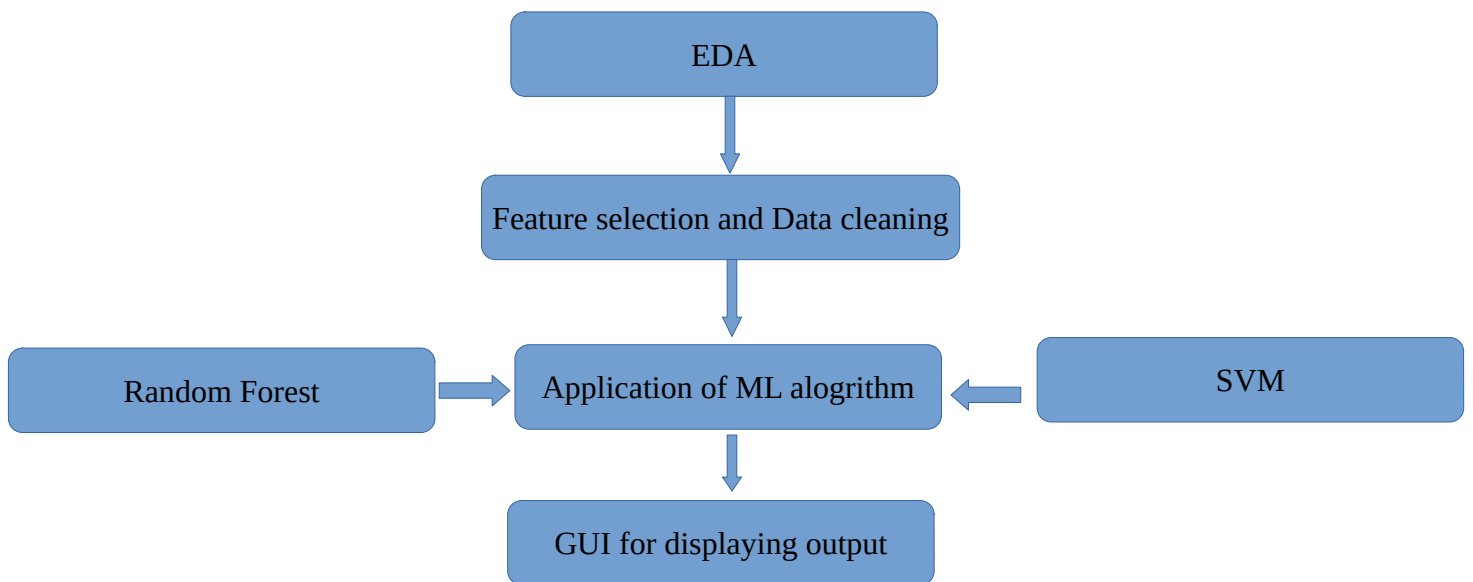
The concept has been development using a Front end Framework such as Angular js and React js.

Business Opportunity.

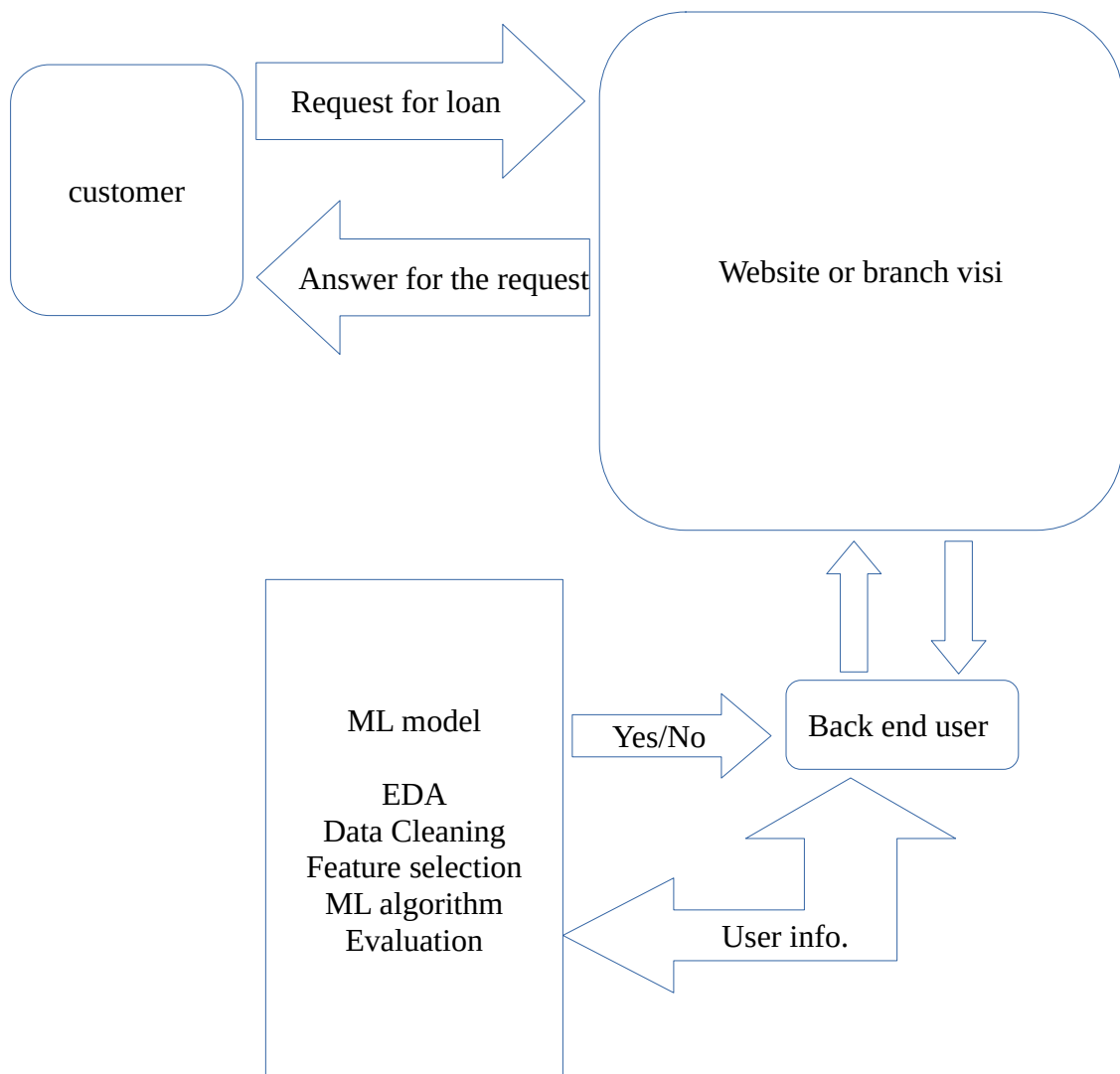
Here the Machine Learning model along with implementing it with small banks and funding clubs can be implemented to other sectors of the business or Banking sectors. In the business such a model can be implemented to categorize to which particular customer an instalment scheme should be given or categorizing the number of customers and find out whether a particular offer for a price is profitable or finding out which are possible flaws in a particular business and creating a more focus on those areas. Along with this this can be implemented in educational institutions in categorising the weak students to put an extra effort on them. The model can also be implemented in finding the factors which lead to downfall of the economy and other sectors.

Final Project Prototype





The above figure is diagram of the proposed model.



The above is a Use Case diagram.

Customer :

He or she is the user who wants the loan for a specific purpose.

Website or Direct branch visit:

It is a means through which the customer will interact with the person in the bank or the back end user. This section also contains blocks or form which are used to collect user details and information.

Back end user:

He/she is the person who actually receives the request and the data from the user. The data is then fed to the machine learning model.

ML model:

This is the machine learning model which is used. Here the steps implemented by the ML model are:

- EDA – Exploratory Data Analysis.
- Data Cleaning – Making the data free from null values and string objects to make it suitable for machine learning models to process.
- Feature selection – Here the important attributes regarding the are considered for performing decision making regarding approval or decline of user request.
- Application of ML algorithm – Here the dataset is splitted into training set, validation set and test set. Then suitable algorithms are applied on them.
- Evaluation – Here evaluation of the output of the model is done using classification reports and confusion matrices.

Then the end results are provided to the back end user and based on them he or she reaches a conclusion of whether or not to provide a loan the person. The end result (either a yes or no) is sent to the user directly or through a website.

Product Details:

•How does it works?

As already explained in the above sections. Here the request is taken from the user along with their information. The data is then fed to the model. The model then analyses the data and based on that predicts whether or not the person will be able to pay back the loan.

•Data Sources

Data Sources are the small business owner, students and people who opt for the loan.

•Software needed

Here we shall need Ms Excel for collections, Jupyter notebooks, python and its libraries for implementing and manipulating the models and also several machine learning libraries.

For end to completion of the system, we also may require a website and for that we may require tools such as Visual Studio Code for building the website and languages such as Html, Css, Javascript and Python and.

•Team Required

1. Web developer for building the website.
2. Data Analyst for analysing the data and cleaning it.
3. Machine Learning expert for proper implementation of the ML model.

• Algorithm

I have implemented the model using the Random Forest. However implementation of SVM is also recommended.

Training the Random Forest model

Now its time to train our model!

Create an instance of the RandomForestClassifier class and fit it to our training data from the previous step.

```
In [29]: from sklearn.ensemble import RandomForestClassifier

In [30]: rfc = RandomForestClassifier(n_estimators=600)

In [31]: rfc.fit(X_train,y_train)

Out[31]: RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
                                max_depth=None, max_features='auto', max_leaf_nodes=None,
                                min_samples_leaf=1, min_samples_split=2,
                                min_weight_fraction_leaf=0.0, n_estimators=600, n_jobs=1,
                                oob_score=False, random_state=None, verbose=0,
                                warm_start=False)
```

Code Implementation

Here is a link whether i have implemented the model.

https://github.com/bipro235/My_Data-science-projects/blob/main/my%20project%20solutions%20.ipynb

Conclusions

Overall the project is a simple project. However there might be few complexities while implementing such a model to the real world. Efficient effort and teamwork can make it a successful one.