

“Prediction of Loan Defaulters”

*(Collaboration of MFI & Telecom Operator to provide mobile balance on credit to “Lower Income Group”)*

Submitted by:

Vinay Kumar

**ACKNOWLEDGMENT**

I would like to thank the FlipRobo team, without whom I would not have been able to complete this project.

**INTRODUCTION**

* Business Problem Framing

A Microfinance Institution (MFI) is an organization that offers financial services to low income populations. Microfinance services (MFS) becomes very useful when targeting especially the unbanked poor families living in remote areas with not much sources of income.

Today, microfinance is widely accepted as a poverty-reduction tool, representing $70 billion in outstanding loans and a global outreach of 200 million clients.

They are working with one such client that is in Telecom Industry from Indonesia. They are a fixed wireless telecommunications network provider. They understand the importance of communication and how it affects a person’s life, thus, focusing on providing their services and products to low income families and poor customers that can help them in the need of hour.

They are collaborating with an MFI to provide micro-credit on mobile balances to be paid back in 5 days. The Consumer is believed to be defaulter if he deviates from the path of paying back the loaned amount within the time duration of 5 days. For the loan amount of 5 (in Indonesian Rupiah), payback amount should be 6 (in Indonesian Rupiah), while, for the loan amount of 10 (in Indonesian Rupiah), the payback amount should be 12 (in Indonesian Rupiah).

In order to improve the selection of customers for the credit, the orgnization wants some predictions that could help them in further investment and improvement in selection of customers.

* Review of Literature

<https://microfinanceinfo.com/micro-financial-institutions/#:~:text=A%20microfinance%20institution%20is%20an,is%20regarded%20as%20microfinance%20institutes>.

* Motivation for the Problem Undertaken

My first objective of this project is to extract the pattern from which our client would find out that this person is can be a defaulter or not means they should give them the loan or not.

**Analytical Problem Framing**

* Mathematical/ Analytical Modeling of the Problem

1. Data Manipulations Using Pandas Python.

2. Univariate analysis

3. Bivariate analysis

4. Model Building

* Data Sources and their formats

Raw Data file in .csv format

* Data Preprocessing Done

1. Dropping rows with –ve values in a DataFrame
2. Changing the Index of a DataFrame
3. Tidying up Fields in the Data
4. Cleaning the Entire Dataset Using the applymap Function
5. Renaming Columns and Skipping Rows
6. Python Data Cleaning: Recap and Resources

* Hardware and Software Requirements and Tools Used

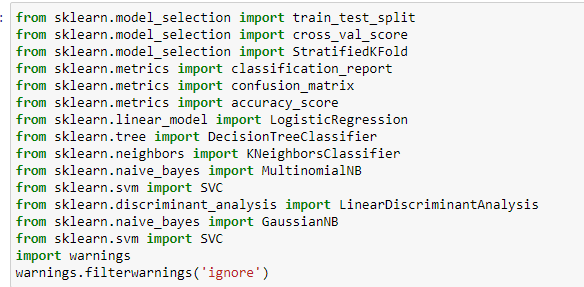
1. Personal Laptop with Intel i3 processor, 4GB DDR3 RAM.
2. Python Jupyter Notepad software
3. MS-Excel

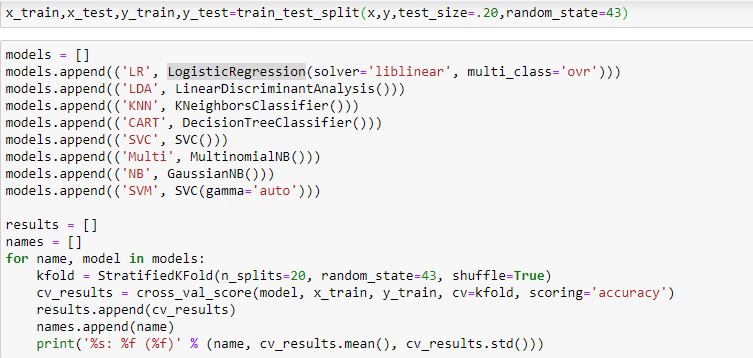
**Model/s Development and Evaluation**

* Testing of Identified Approaches (Algorithms)

1. Logistic Regression
2. DecisionTreeClassifier
3. KNeighborsClassifier
4. MultinomialNB
5. LinearDiscriminantAnalysis
6. GaussianNB
7. SVC

* Run and Evaluate selected models





* Key Metrics for success in solving problem under consideration

1. Cross Validation Score.
2. accuracy\_score
3. confusion\_matrix
4. classification\_report

* Interpretation of the Results

After all processing & analysis, i found that Logistic Regression is the best suitable algorithm for this problem. Model accepts more than 85% respectively.

**CONCLUSION**

1. Key Findings and Conclusions of the Study

I observed that data cleaning is the most important part to under the clear picture of any data set. Afterwards we will find out logical pattern to understand & solve the business problem respectively.

1. Learning Outcomes of the Study in respect of Data Science

After all processing & analysis, i found that Logistic Regression is the best suitable algorithm for this problem. Model accepts more than 85% respectively.