International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IJIEMR Transactions, online available on 4th Aug 2021. Link :http://www.ijiemr.org/downloads.php?vol=Volume-10&issue=ISSUE-08

DOI: 10.48047/IJIEMR/V10/I08/06

Title **CALCULATION OF LOAN DEFAULTER BASED ON LOAN REPAYMENT HISTORY**

Volume 10, Issue 08, Pages: 35-40

Paper Authors
Miss. YT RAJYALAKSHMI, Dr. E.SREEDEVI

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

CALCULATION OF LOAN DEFAULTER BASED ON LOAN REPAYMENT HISTORY

Miss. YT RAJYALAKSHMI, IVth semester, Student In Dept. Of MCA, SVIM,India,Email Id:ytrajyalakshmi@gmail.com

Dr. E.SREEDEVI, Assistant professor, dept of MCA , Sree Vidyanikethan Institute Of Management from SV University, A.P, INDIA

ABSTRACT

With the improvement in the banking sector, more individuals are requesting for bank loans. However, banks have limited assets and can only issue loans to a limited number of people, so determining who the loan can be provided to and who would be a safer option for the bank is a common procedure. So, in this article, we attempt to decrease the risk factor associated with picking the safe individual in order to save the bank time and money. This is accomplished by mining Big Data of past records of persons to whom the loan was previously issued, and the computer was taught using a machine learning model based on these records/experiences to provide the most accurate result. The major goal of this article is to determine whether or not allocating a loan to a certain individual is safe. There are four sections to this study. I Data collection (ii) Machine learning model comparison using acquired data (iii) System training using the most promising model (iv) Testing

Keywords: ***financial inclusion; credit score; big data; machine learning; airtime***

I INTRODUCTION

Almost every bank's primary activity is the distribution of loans. The profit gained from the loans distributed by the banks accounts for the majority of the bank's assets. In a banking setting, the primary goal is to place their funds in safe hands. Many banks and financial firms now grant loans after a lengthy process of verification and validation, but there is no guarantee that the chosen candidate is the most worthy of all applicants. We can anticipate if a given applicant is safe or not using this method, and the entire process of feature validation is automated using machine learning techniques. The downside of this model is that it gives various weights to each element, but in reality, a loan might sometimes be authorised only on the basis of a single strong factor, which is not feasible with this method. Loan Prediction is extremely beneficial to both bank employees and applicants. The purpose of this paper is to give a quick, straightforward, and efficient method of selecting qualified applicants. It may give the bank with unique benefits. The Loan Prediction System can compute the weight of each feature involved in loan processing automatically, and the same characteristics are processed on new test data in accordance with their associated weight. The applicant might be given a deadline to determine whether or not his or her loan will be approved. The Loan Prediction System allows you to skip to a specific application and review it on a priority basis. This paper is intended just for the management of a bank or finance business; the whole prediction process is conducted in private, and no stakeholders will be able to influence the outcome. The results for a specific Loan Id can be sent to other bank departments, allowing them to take appropriate action on the application. This facilitates the completion of additional formalities by all other departments.

II RELATED WORK

The payback time of a loan is influenced by the borrower's liquidity situation as well as the investment's economic life. Repayment schedules must be flexible enough to be changed according to the borrower's cash flow pattern. Banks also employ certain relevant lending principles as guiding principles in addition to these credit policy instruments (Zena 2009). Borrowers' perceived need, competency or repayment capacity, and personal character are among them.

According to William (2007), no logical lender wants or will hand over funds to a borrower to run and invest in a business or endeavour in which the borrower has no or limited experience. From the lender's and borrower's perspectives, this requirement for successful borrowing should be obvious. Lenders must be more confident that the individual or persons borrowing the cash have the necessary knowledge and skills to manage the funds and that the business is done responsibly on a daily basis. This is required to ensure that the firm produces good outcomes and that the loan is repaid with interest and on time. Once again, the issue is one of risk. From the bank's perspective, the lesser the risk of loan, the more expertise and talent the borrower has demonstrated in the past.

As the firm works from period to period, cash flow indicates the cash income flowing into the business and the funds going out in paid costs. On a month-to-month basis, the net outcome is either negative or positive. The lender is looking for positive cash flows that are at least sufficient to repay the loan. The Income Statement provides a summary of the company's total income during the previous year, as well as a breakdown of all of the company's costs. When these two numbers, one positive and one negative, are added together, the net profit of the company activities for the time under consideration is calculated. The Balance Sheet may be regarded of as a snapshot of the company's financial "health" at a specific point in time, usually towards the end of the year. The balance sheet shows the assets (positive values of the company's assets) and liabilities (negative values of the company's liabilities) (the negative obligations that the business owes and is obligated to pay in the short or long run). Liabilities in addition to the owner's or shareholder’s equity (value of the ownership) will equal (balance) the value of the assets.

Successful loan repayment is defined as the capacity to repay the loan according to the loan agreement, whereas loan defaulting is defined as the inability to return the loan due to failure to finish the loan according to the loan agreement or neglecting to service the loan. She discovered a substantial link between major sources of income, diversion of cash, domestic difficulties, and loan defaulting in her study on the reasons of default in government micro credit programmes in Kenya. Reza and Mansoori (2008) studied the factors impacting the repayment behaviour of farmers in Iran's Khorasan-Razavi area who obtained loans from agricultural banks. In the year 2008, They utilise a logit model to discover the characteristics that influence loan repayment performance in their study approach. Farmers' delayed repayment of loan instalments to the bank is characterised as a dependent variable. As a result, the value of the dependent variable will be one if the farmer has not made any late payments, and 0 otherwise. The study was conducted by conducting a survey and filling out questionnaires for 175 farmers in rural areas of Iran's Khorasan-Razavi province.

III. SYSTEM ANALYSIS

EXISTING SYSTEM

With the improvement in the banking sector, more individuals are requesting for bank loans. However, banks have limited assets and can only issue loans to a limited number of people, so determining who the loan can be provided to and who would be a safer option for the bank is a common procedure.

PROPOSED SYSTEM

We aim to decrease the risk factor in selecting the safe individual in the suggested system in order to save the bank a lot of time and money. This is accomplished by mining Big Data of past records of persons to whom the loan was previously issued, and the computer was taught using a machine learning model based on these records/experiences to provide the most accurate result. The major goal of this article is to determine whether or not allocating a loan to a certain individual is safe. There are four sections to this study. I Data collection (ii) Machine learning model comparison using acquired data (iii) System training using the most promising model (iv) Testing

IV.METHEDOLOGY AND MODELS

The training data set is now sent to the machine learning model, and the model is trained using this data set. Every new applicant's information entered on the application form serves as a test data set. Following the testing procedure, the model predicts if the new applicant is a good candidate for loan approval based on the inferences drawn from the training data sets.For android application prediction, six machine learning classification models were utilised.

The R open source programme is used to create the models. R is released under the GNU General Public License. The following sections provide a summary of each model.

The importance of nonlinearity and the possible benefits of employing an ensemble method are investigated using three machine learning models. It is thus feasible to find the best effective model structure for explaining the links between the explanatory factors and default for airtime lending by combining the three models. First, logistic regression (LR) uses linear connections to generate binary classifications (Cox 1958). LR has long been utilised as a basic model structure that serves as a baseline for assessing classifier performance. Second, a decision tree (DT) is built using a nonlinear model to assess the possibility for improvement (William 1959). While DT provides the potential of a nonlinear model structure, it also provides a set of principles that are very simple to apply. Third, by averaging over a number of decision trees, an ensemble technique known as Random Forest (RF) is used (Breiman 2011). It is feasible to determine if ensemble approaches may provide any improvements beyond the LR and DT techniques by incorporating the RF model.

Decision Tree

All characteristics or features must be discretized according to the decision tree's fundamental algorithm [7]. Feature selection is dependent on which features provide the most information. IF-THEN rules can be used to express the knowledge shown in a decision tree. Quinlan's C4.5 classification methods are used into this model.

Random Forest

Random forests [8] are a characterisation (and relapse) group learning system that works by constructing a large number of Decision trees over time and producing the class that is the mode of the classes yielded by individual trees.

SVM

Support vector machines are learning models that employ the association r learning method to evaluate data and identify pattern knowledge, which is then used to classify applications. SVM may use the kernel technique to produce a regression that verifies the mapping of their inputs into high dimensional feature spaces.

Data Analysis

The next stage in developing a credit scoring model is to conduct an exploratory study of the data and offer summary statistics about the variables, based on the findings of this meta-analysis. The research period is dependent on the data provided by ComzAfrica. Data for the period 1 January 2016 to 30 June 2017 was acquired after negotiating legal and confidentially agreements with Comz Africa as part of the CMU-Africa practicum. These records were linked to three million loans made by 46 thousand clients. Given the commercial sensitivity of the investigation, more time was necessary to get approval for publishing of the data. Loans are examined from 1 January 2016 to 30 April 2017 in order to assess each customer's performance in the three months after the prior loan (performance window). Defaulters are those who have not paid within the performance timeframe.

Banking Service

The Development Bank of Ethiopia's banking services are primarily aimed at its clients, and they were established with the primary goal of supporting customers' commercial operations and loan services.

As a result, the bank offers a variety of ancillary services in both domestic and international banking. Import and export L/C, as well as settlement of intangible commodities such as consultation, installation, and commissioning services, are among the services given to its clients in the area of international banking operations. The domestic banking wing, on the other hand, provides current accounts and inter-bank check clearance in Addis Ababa city to its consumers in addition to servicing the bank's staff. Domestic banking also helps and supports the bank's disbursement and loan collection efforts in general, although deposit mobilisation isn't a priority because the bank isn't required to do so.

CONCLUSION

It may be confidently inferred that the result is a highly efficient component based on a good examination of positive points and restrictions on the component. This application is operational and complies with all Banker criteria. This component can easily be integrated into a variety of different systems. There have been several instances of computer malfunctions, content mistakes, and, most importantly, the weight of characteristics in automated prediction systems that have been rectified. As a result, the so-called software might be improved in the near future to make it more safe, dependable, and dynamic in weight adjustment. This prediction module can be integrated with the automated processing system module in the near future. The system is trained on an old training dataset; but, in the future, software can be developed such that new testing data can be included in training data after a predetermined period of time.

REFERENCES

- Aker, Jenny C., and Isaac M. Mbiti. 2010. Mobile Phones and Economic Development in Africa. Journal of EconomicPerspectives 24: 207–32. [CrossRef]
- Baklouti, Ibtissem. 2013. Determinants of Microcredit Repayment: The Case of Tunisian Microfinance Bank.In African Development Review 25: 370–82. [CrossRef]
- Belson, William A. 1959. Matching and Prediction on the Principle of Biological Classification. Journal of the RoyalStatistical Society. Series C (Applied Statistics) 8: 65–75. [CrossRef]
- Blanco, Antonio, Rafael Pino-Mejías, Juan Lara, and Salvador Rayo. 2013. Credit scoring models for the microfinance industry using neural networks: Evidence from Peru. Expert Systems with Applications 40: 356–64.[CrossRef]Breiman, Leo. 2001. Random Forests. Machine Learning 5: 5–32. [CrossRef]
- Chen, Gregory C., and Pedro Xavier Faz De Los Santos. 2015. The Potential of Digital Data: How Far Can It Advance
- Financial Inclusion? Focus Note 100. Washington, DC: CGAP.
- ComzAfrica. 2017. Products|ComzAfrica. Available online: http://www.ComzAfricafrica.com/

products(accessed on 11 October 2017).

- Cox, David R. 1958. The Regression Analysis of Binary Sequences. Journal of the Royal Statistical Society: Series B(Methodological) 20: 215–42.
- Cubiles-De-La-Vega, María-Dolores, Antonio Blanco-Oliver, Rafael Pino-Mejías, and Juan Lara-Rubio. 2013.Improving the management of microfinanc institutions by using credit scoring models based on StatisticalLearning techniques. Expert Systems with Applications 40: 6910–17. [CrossRef]
- Deloitte. 2017. Credit scoring Case study in Data Analytics. Available online: https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Financial-Services/gx-be-aers-fsi-creditscoring.pdf (accessed on12 October 2017).
- Madise, Sunduzwayo. 2015. Mobile Money and Airtime: Emerging Forms of Money. SSRN Electronic Journal.
- [CrossRef]J. Risk Financial Manag. 2020, 13, 180 11 of 11Mode. 2017. Products. Available online: http://www.mo-de.com/products/ (accessed on 11 October 2017).
- Pell, Nicholas. 2017. A Secret History of Credit Scores: Who Determined What Matters and Why. Available online:https://www.thestreet.com/personal-finance/credit-cards/a-secret-history-of-credit-scores-who-determined-what-matters-and-why-13097739 (accessed on 12 October 2017).
- Safaricom. 2017. OkoaJahazi. Available online: https://www.safaricom.co.ke/personal/plans/gettingstarted/okoa-jahazi (accessed on 11 October 2017).
- Tang, Yuchun, Yan-Qing Zhang, Nitesh V. Chawla, and Sven Krasser. 2009. SVMs Modeling for Highly Imbalanced