2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI) | 978-1-7281-9544-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/SACI51354.2021.9465561

Using Machine Learning Algorithms to create a Credit Scoring Model for mobile money users

Monica Charles Mhina
Electrical and Computer Engineering Department
McGill University
Montreal, QC, Canada
monica.mhina@mail.mcgill.ca

Fabrice Labeau
Electrical and Computer Engineering Department
McGill University
Montreal, QC, Canada
fabrice.labeau@mcgill.ca

Abstract— **Statistical and artificial intelligence methods are used extensively to analyze credit and evaluate the credit risk of loan application clients. In this paper, mobile transaction data of agents from a digital payment switching company that is interested in offering microloans to its agents were used to determine the agent's creditworthiness. Traditional credit scoring methods do not work for these agents as the transaction data is not recorded by a full-service financial institution like a bank. Different data manipulation techniques were explored to present data into features that can be used for scoring. The effects of resulting features were explored using correlation and singular value decomposition, and clustered using *K*- means clustering to assess creditworthiness. After clustering agents into *K* groups, these groups were clustered again to determine low-risk agents, and a formula to determine how much credit can be extended to low-risk agents was devised.**

Keywords—mobile payment; feature selection; credit scoring model

I. Introduction

Credit is based on trust, trust that a lender can lend money or resources to a borrower, and the borrower returns in a given amount of time. Credit assessment is essential for the financial success of both lenders and borrowers. Lenders, usually financial institutions, collect information from borrowers to be able to determine their creditworthiness. A credit score is a score assigned by a scoring model that tells a financial institution or a lender the risk of lending to either an individual/customer or a company [1].

There are different types of credit scoring. These are fraud detection, collection scoring, behavioral scoring, and application scoring. Fraud detection scoring deals with detecting fraudulent applications; collection scoring deals with categorizing insolvency levels; behavioral scoring deals with assessing the risk of lending to existing customers; and application scoring deals with assessing the risk of lending to new customers [1]. For this project, the focus will be on application scoring, which is usually referred to as credit scoring.

The two main credit scoring models' developers are FICO and VantageScore [2]. These credit scoring companies design scoring models and software that most financial institutions in the world use. FICO and VantageScore are considered traditional scoring methods that use a mix of customer's payment history, amounts owed, length of credit history, new/recent credit, utilization of available credit, and a mix of all credits to determine a customer's credit score [4] [5]. Therefore, customers who are new to the credit market, have no open credit accounts, don't use their credit frequently, or have any recent credit activity are considered unscorable. According to VantageScore, unscorable customers are both a cost and a growth opportunity for lenders. Lenders have to work with credit scoring model developers like VantageScore to develop models that go beyond traditional scoring methods [3].

A. Background/Motivation

According to the World Bank Findex survey in 2017, financial inclusion in Sub-Saharan Africa has "increased dramatically" from 23% in 2011 to 43% in 2017 [6]. This more than almost 50% increase can be attributed to mobile money accounts or the so-called new digital finance space. Mobile money is a service that stores funds in a secure electronic account, linked to a mobile phone number. Mobile money is often provided by the same companies that run the country's mobile phone services, often referred to as mobile network operators (MNO's). Mobile money is a popular alternative to both cash and banks because it's easy to use, secure, and can be used anywhere there is a mobile phone signal [7].

While the number of bank accounts or mobile money accounts is a good measure of financial inclusion, it is important to assess whether such inclusion is not only reaching but also helping the common beneficiaries. In Tanzania specifically, 56% of adult population in 2017 was financially included via a registered account with a full-service financial institution. As with the rest of Sub-Saharan Africa, this financial inclusion is driven by mobile money whereby 55% of adults had mobile money accounts, versus 9% who had separate bank accounts, and 3% who had NBFI (nonbank financial institution) accounts [8].

The increase in financial inclusion attributed to mobile money has increased digital transactions via mobile money. However, these intermediate transactions, whether paid for by cash or mobile money, are not considered in the formal credit scoring system. In emerging economies, credit scoring remains a challenge as a significant portion of the population has a bad (if any) or no credit score due to doing digital transactions by mobile money as opposed to a registered account with a full-service financial institution like a bank. This is a gap in financial inclusion that needs to be closed.

This project stems from the need to increase visibility and add value to users who do not have a formal credit score in the financial world. This way of credit scoring is called 'alternative credit scoring' and can be done by studying and drawing knowledge from behavioral trends from online presence, psychometric testing, non-financial payment streams, unbanked transactions, etc. [9]. This data can be analyzed, and an alternative credit score can be computed for mobile money users.

For this paper specifically, digital transactions done using cash or mobile money or bank cards that go through a transaction switching company will be the focus. A customer usually approaches a transaction switching company agent to pay for goods and services such as water or electricity bills. An agent who helps customers perform these transactions would not have these on record with a full-service financial institution but rather with the transaction switching company. These agents need alternative credit scoring and are the focus of this project.

The agent usually needs to have e-money (also called float) to be able to conduct transactions on behalf of the customer and earns a percentage of transacted value as a commission, which is deposited into their mobile money accounts. The data on transactions done by agents are readily available, and since agents need to have e-money to do transactions, the ability to be able to borrow from time to time will alleviate moments when a customer wants services, but the agent has no e-money. Also, having access to credit will incentivize agents to do more transactions and keep a track record to qualify for more credit.

Other factors such as agents' behavioral trends from online presence were ignored. This is not only because the data is not readily available but also because of privacy issues and misclassification [2]. Learning customer behavior through their online trail has emerged as an essential part of big-data profiling. This has raised concern in user information protection as well as the risk of inaccurate information from online trails [2]. In addition, this project looks into providing microloans to the agents, so if traditional features such as the length of credit history and types of credit considered by FICO or VantageScore were used, they might be too extreme and not apply to these agents, deeming them unscorable by traditional scoring standards.

The rest of the paper is organized as follows: Section II covers related works. Section III describes details of the proposed prediction model. Section IV discusses experimental results while Section V provides discussion of results, concluding remarks and future work.

II. Related Works

To inform their lending decisions, financial institutions need to predict loan default probability of the borrower [26]. The notion of credit scoring involves assigning a quantitative score that defines good and bad borrowers and has been developing over the years, as summarized by Baesans *et al.* [10]. In their ten-year update of Baesans *et al.,* Lessmann *et al.* explored 41 classifiers, including K- means that have been used for credit scoring [11]. Ha *et al.* analyzed the importance of feature selection before computing the credit score and concluded that feature selection saves time (runtime and evaluation personnel time) and increases the accuracy of the classifier [27]. Additionally, Addo *et al.* suggest re-checking model performance on selected features [28].

There are two types of risk models, corporate and retail. While corporate risk models use data from balance sheets etc., retail risk models use transactional data from the customer history, etc. Although the traditional banking industry-standard methods for credit risk assessment are linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and logistic regression (LR); Lasemann *et al.* concluded that heterogeneous ensemble classifiers such a K- means had been proved to do well in credit scoring [11]. Nan-Chen *et al.* use K- means and a neural network to design their hybrid credit scoring model [12]. Hsieh *et al.* use K- means to pre-process their data in what they call class-wise classification, which enables an efficient ensemble classifier [13].

Sadatrasoul *et al.* classify credit scoring models into three categories, single classifier, ensemble, and hybrid. Single classifiers such as Support Vector Machines (SVM) and decision trees group applicants as good or bad. Ensemble methods aggregate multiple classifiers to improve performance accuracy. Hybrid methods combining clustering and classification involve two or more models whose combination covers their individual weaknesses [1].

Outliers are data that lie outside the normal range; in other words, they are extreme values. Yu *et al.* discuss three different types of outliers: collective, contextual, and global. These significantly deviate from the whole data, ones that significantly deviate from other data points in a given context, and ones that significantly deviate from other data points, respectively [14]. Yu *et al.* argue that outliers affect K- means clustering by destabilizing cluster centers since means are sensitive to extreme values. When centers are destabilized, the algorithm's efficiency decreases; hence outliers must be eliminated [14]. Agents who are global outliers were removed from final calculations.

Blanco *et al.* designed a credit scoring model for microfinance using the multilayer perceptron approach (MLP) [15]. Shema designed a credit scoring model that uses airtime recharge data using random forest classifiers to predict default probability [16]. Beyond machine learning techniques, Mosavi *et al.* explored deep learning models such as binary classification techniques to perform credit risk analysis and loan pricing [26]. Despite much research, to the best of our knowledge, the credit scoring literature does not include informal financial services companies who are interested in short-term digital microloans to their employees. Hence, a hybrid method of clustering + clustering will be explored to score our agents.

III. Methods

The system in Figure 1 was designed to create a credit scoring model for agents.

A. Data

Goods and service providers such as utility companies usually have a collection account with a transaction switching company connected to their bank account. The account with a transaction switching company alleviates the pain of having an account with every mobile money provider. For example, there are x official mobile money providers in a country. Hence, a utility company would need to have an account at all x companies to receive payments from their customers. A unified account is more convenient. The transaction switching company has agents all over the country who assist goods and services providers in the selling of their goods and services.

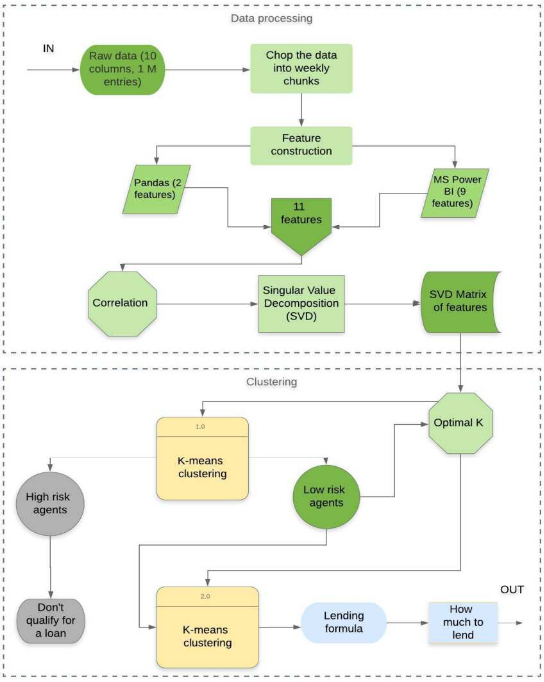


Fig. 1. System design

These agents are equipped with three possible tools to use to conduct transactions. These are the point-of-sale terminals (POS), mobile phone (via USSD (Unstructured Supplementary Service Data) codes or an application), or a web browser. The cash flow from a customer to a goods and service provider's bank account via digital transaction switch company is shown in Figure 2 below.

When an agent helps a customer to pay for goods and services, information from the transaction is recorded by the transaction switching company. This information includes agent ID, transaction type, transaction ID/reference number, utility code/reference number, and amount transacted.

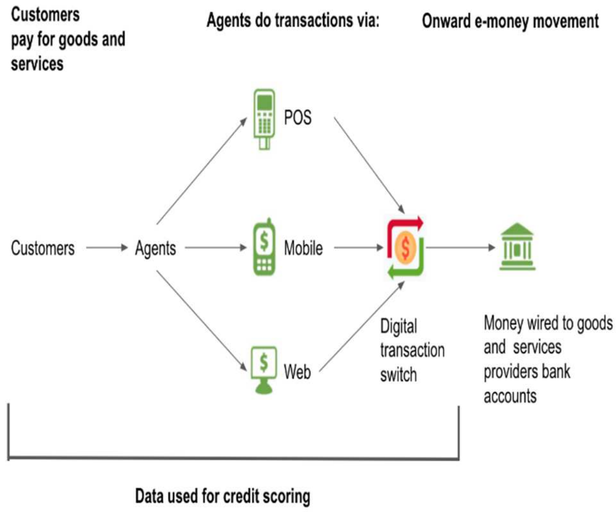
The project will be based on a large data sample provided by the transaction switching company, which contains more than 1 million transactions done by about 6,000 agents in Tanzania for one month. The information will be processed in the system to create features that will be used in our clustering + clustering approach.

B. Feature Construction

A total of 11 features were constructed. These are amount transacted (total, maximum, minimum, average), total number of transactions, agent (revenue, float (e-money), float top-up), number of unique services, standard deviation (number of transactions, amount transacted).

These 11 features were selected, keeping in mind some of the most common ways an agent can use to cheat the credit scoring system. An agent can attempt to boost their volume of transactions by doing low-value transactions, but this will end up hurting their score as the total, minimum, maximum, and average value of the amount transacted are all considered. If an agent targets just a particular good or service, this will be captured in the number of unique services. Also, if an agent transacts only one large amount every now and then, the number of transactions, as well as the spread of those transactions, will be captured. The selected features captured these and many other scenarios and were deemed enough to score agents. These 11 features were extracted for all four weeks observed in this project, bringing the total number of features to 44.

Fig. 2. Data acquisition channel (* icons courtesy of [25])



C. Feature Selection

After constructing the features, it is important to determine whether they provide different information or not and rank their importance. This is important because credit data contains redundant data that needs to be filtered by feature selection to increase the accuracy of the classifier [26]. To do so, the correlation and singular value decomposition (SVD) were employed to obtain independent features and break down the feature matrix into singular values to obtain unique dominating features, respectively.

1) Correlation

Correlation is a statistical method of determining linear dependence between variables (in this case, features). Correlation aims to reduce dimensions by dropping one of two highly dependent variables. Pearson correlation [17] was used on our constructed features. If two features were found to be highly correlated, one of them was dropped as a significant amount of information they provide is already provided by the other. Overall, four features were dropped, bringing the total number of features for all four weeks down to 40 features.

2) Singular Value Decomposition (SVD)

Singular value decomposition is a statistical method used to reduce dimensions of features in Machine Learning. Singular value decomposition also extracts dominating features in a large database [18]. After decomposing the data matrix, the order of magnitude of the singular values ranged from 10^1 to 10^9. Only features with values with magnitude greater than 10^5 are desirable because upon graphing the singular values, the values plateau after the 10^5 point. This brought down our total number of features from 40 that we had after correlation to 24.

3) K-Means Clustering

Clustering is a way of grouping observations such that observations that are similar to each other are in the same group or cluster [20]. K-means clustering is one of the most common Machine Learning algorithms used mainly for unsupervised learning. K-means clustering was chosen specifically because there is no ground truth data to be used for scoring [1], and because it has the highest accuracy of an unsupervised learning algorithm for credit scoring [27]. K- means clustering aims to partition observation into K clusters and assign each observation (in this case, agent) to a cluster with the nearest mean/centers/centroid. The number of clusters, K, has to be specified before K- means clusters the data. With each iteration, the centroids are updated/iterated while optimizing the Euclidean distance from one centroid to another. The algorithm stops when the error does not decrease significantly [21]. The cost function sums the distortions of the clusters and can be calculated using the following formula [22]:

$$J = \sum_{k=1}^{K} \sum_{i \epsilon C_k} \|x_i - \mu_k\|^2$$

where J is the cost function, x_i is the feature vector for agent i, μ_k is the centroid of cluster k, and C_k is the set of agents currently assigned to cluster k. Determining the optimal value for K is very important. One heuristic way of estimating the optimal K is the elbow method [20].

IV. Results

The clustering + clustering technique was used to score agents. This means agents were clustered into groups with similar performing agents. Then, clustered agents in the same group were clustered again to further differentiate the agents into a smaller group of most trustworthy agents. This is a combination of two clustering techniques [1] whereby the first clustering can reduce the data; hence it can be considered as a data reduction step in addition to correlation and singular value decomposition. For our first clusters, the elbow is forming at about K= 3. Hence, the optimal value of K was estimated to be 3 as seen in Fig. 3 below.

The optimal value of K was used to compute K- means clustering for our reduced dimension data matrix A'. The K- means labels were then appended to our data matrix with all 44 features constructed in section III. Despite being a very common algorithm, K- means clustering is very sensitive to noise and outliers [9]. Hence, removing outliers is one of the most effective ways to improve the accuracy of K- means clustering. Some agents were observed to be outliers because no matter the value of K, they always ended up in their own group by themselves. These outliers were dropped from features data, and dimensionality reduction (correlation and singular value decomposition) was performed again before recomputing K- means clustering.

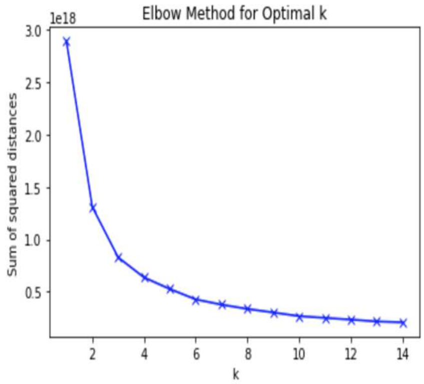


Fig. 3. Elbow method for the optimal value of K

After removing outliers, to observe the characteristics of each cluster in terms of features, multiple features were graphed against each other. Agent float top-up is an important feature for both agents and the transaction switching company since an agent cannot help customers pay for goods and services without having float (e-money). Also, the weekly loan that agents will receive will be in float. The relationship between agent float top-ups to other features was explored. It was observed that agents with the lowest total amount transacted have the most float top-ups.

Agents with the lowest standard deviation in the total amount transacted have the most float top-ups, which demonstrates that our agents are transacting evenly across the week. If an agent is transacting evenly across the week, they will need float; hence their float top-ups being high is expected.

Agents with the lowest float have the greatest number of float top-ups. Low float means that if an agent has to do more transactions, they will need to get more float; hence, many float top-ups are expected. The distinction was not clear enough for the maximum amount, minimum amount, average amount, total number of transactions, unique services, and agent revenue.

In addition to float top-ups, another important feature to both agents and the transaction switching company is agent float. It was observed that the higher the agent float, the lower the standard deviation in the total number of transactions. This means that the more float an agent has, the more evenly distributed are the number of transactions across the week. This is expected because when an agent has more float, they can conduct transactions at any time.

The last important feature to both agents and the transaction switching company explored is agent revenue/commission. Agent revenue is not affected by the total amount as well as the standard deviation of the amount transacted as in 0. This was not expected since the more money you transact, the more you should earn. This suggests that an agent's commission is not necessarily dependent on the total amount they transact but on other factors as well.

The characteristics distinction between agent revenue and each of the following features (i.e., maximum amount, minimum amount, average amount, total transactions, the standard deviation of the number of transactions, unique services, agent float, and agent float top-ups) was not clear enough when graphed.

Overall, agent float top-ups, agent float, and agent revenue were chosen as primary features to be used to observe characteristics of our three groups. In summary, float top-up has an inverse relationship to the total amount, agent float, and standard deviation of the transacted amount. Agent float also has an inverse relationship to total transactions and the standard deviation of the number of transactions. Lastly, agent revenue has no relationship to the total amount and standard deviation of the total transacted amount.

Our 6,046 agents are now divided into three groups: 0, 1, and 2, with 5,946, 12, and 88 agents respectively. Group 0 includes a range of agents from those who transacted one week out of the four, to those who transacted in all four weeks. Group 1 and 2 have agents who transacted in all four weeks. Since no clear distinction was established among the groups, all three groups were analyzed to be clustered again to refine our agents further. The 5,946 agents in group 0 were to be clustered again. Before clustering, singular value decomposition was applied and the final number of features to be used for clustering was 24. Upon using the elbow method to determine the optimal value of □, the elbow was estimated to be at □ = 3. Hence, clustering was performed for □ = 3. The 5,946 agents were clustered into three groups, 0,1 and 2, with 5,382, 127, 437 agents, respectively. Multiple features were graphed against each other to observe the characteristics of each cluster in terms of features. Once again, it was observed that agents with the lowest total amount transacted and standard deviation in the total amount transacted have the most float top-ups. Also, agents with the lowest float had the greatest number of float top-ups, and agent revenue is not affected by the total amount as well as the standard deviation of the amount transacted. It is important to note that all feature interactions explored for group 0 behaved the same way as the feature interactions in the first round of clustering. Inspecting the characteristics of Group 1, no clear elbow could be determined; hence clustering was not performed. Regarding Group 2, since this is a small group of just 88 agents, the second round of clustering was not necessary.

A. Who is a good agent?

Overall, the 12 agents in group 1 were not clustered again due to lack of optimal K. The 88 agents in group 2 were clustered again for $K = 2$ into 21 agents who made a revenue in all weeks and 67 agents who had revenue gaps in a week or more, however, there was no distinction between the two groups. The 5946 agents in group 0 include agents who did not transact in some weeks to those who transacted in all 4 weeks. To further differentiate agents in group 0, the 5946 agents were clustered again for $K = 3$. The new groups had 5382, 127, 437 agents, respectively with similar patterns (the larger group contains agents with varying characteristics). If we base risk behavior on consistency, we could say that agents in groups 1 and 2, as well as some agents in group 0, present the least risk of lending.

However, we could have concluded that agents who consistently transact in all four weeks, like those in groups 1, 2, and some in group 0, are low risk even without K- means clustering. Hence, we need to investigate why most of our agents are grouped into group 0 despite having varying characteristics. The time factor was suspected to be the issue since we merged all four weeks to form 44 features before analyzing our agents. To investigate if time is the issue, agents who transacted in the first week (week 1) were used. As we did previously, 11 features were constructed, and both correlation and singular value decomposition were applied. The reduced features matrix was used for clustering. The optimal value for K was found to be 3. The 3,514 agents who transacted in the first week were clustered into three groups of 3,347, 19, and 148 agents, respectively. Feature characteristics observed for week 1 were the same as those observed in our first round of K- means clustering. The three groups in week 1 were also consistent in a way that most agents are clustered into the first group, and the remaining two groups have a few agents each. Hence, there is not enough evidence to rule out the time factor as the reason why some consistent agents end up in group 0.

Since there is consistency in some agents who transact all four weeks ending up in the dominant group (like group 0) with agents with varying characteristics, to the extent of this project, we can classify group 0 as riskier than group 1 and 2. For those agents who end up in group 0 despite having transacted in all 4 weeks, there needs to be another measure to evaluate them. Perhaps incorporating more features so they can be distinguished from the least consistent agents. Another option that can be explored is to try using data from another month and see if the trend persists. Lastly, we can use the agents in groups 1 and 2 as ground truth data and explore supervised learning methods for agents in group 0.

B. How much to lend?

After determining agents worth lending to, it is important to determine how much the transaction switching company can lend to each agent. Every agent is paid a commission (agent revenue) by the transaction switching company (TSC) to their mobile money accounts for assisting customers purchase goods and services. The agents incur a 5% withholding tax (WHT) on their revenue collected by the TSC on behalf of the government.

The above charges were taken into account when deciding how much money can be lent to an agent. Given that the annual interest rate is 21% and the maximum amount the transaction switching company is willing to lend is 500 USD (1,150,000 Tanzanian shillings), the following formula was devised to determine the maximum weekly loan interest that the company can collect from the agent. This maximum weekly loan interest can eventually be used to calculate how much to lend weekly.

$$\begin{aligned} &Maximum\ weekly\ loan\ \ interest \\ &= \frac{(Agent\ revenue\ \ - 5\%\ WHT\ on\ Agent\ revenue\ )}{(100\% + Weekly\ interest\ rate\ )} \end{aligned}$$

The maximum weekly loan interest formula is based on the fact that the transaction switching company is interested in recovering both their loan and the interest rate. The formula tells the transaction switching company to use the average agent revenue over the past few weeks as the basis for how much to lend. A 21% yearly interest rate is about 0.4% weekly interest rate. So, the maximum interest on the weekly loan should not exceed the average agent revenue. This is because the company can recover the principle when an agent buys more float (e-money) and need to recover their interest from the agent's revenue to avoid agents who try to cheat the system and avoid the interest. Below is an example of how much credit line can be extended to an agent. Let's assume that our agent makes 50,000 Tanzanian shillings a week. Then, this agent can receive a loan whose interest rate is not more than the 50,000 Tanzanian shillings a week minus taxes and service charge.

$$Maximum\ weekly\ loan\ \ interest = \frac{(50{,}000 - 5\%(50{,}000)\ )}{(100\% + 0.4\%\ )} \cong 46{,}688 \text{ Tanzanian shillings}$$

At a 0.4% weekly interest rate, the maximum credit that can be extended to this agent is:

$$Maximum\ credit\ line = \frac{46{,}688*100\%}{0.4\%} \cong 11.67M \text{ Tanzanian shillings}$$

The maximum weekly loan interest is the expected agent revenue for the coming week. However, as a company, we do not want to take away all agent's revenue as interest. For starters, we can decide to take only 5%, which is the same as the government's withholding tax rate. This will bring the maximum credit line down to 583,600 Tanzanian shillings.

$$Reduced\ maximum\ credit\ line = \frac{5\%(46{,}688)*100\%}{0.4\%} \cong 583{,}600 \text{ Tanzanian shillings}$$

The reduced maximum credit line is well within the 1,150,000 Tanzanian shillings (500 USD) that the transaction switching company is willing to lend. Trial and error will be needed during the testing phase to determine the optimal percentage of expected agent revenue that can be taken as interest, and the optimal number of weeks to use to find the average agent revenue.

V. Discussion and Conclusion

Credit scoring systems are common, but the application of credit scoring on transactional data of agents who bank with mobile money is a relatively recent issue. In this report, credit scoring agents whose revenue doesn't go to a full-service financial institution (e.g., bank) account was explored. A method to effectively score these agents was devised. Raw data from the employer (transaction switching company) was used to construct features that were reduced and used to cluster agents using K- means clustering. Also, a formula on how much credit to extend to creditworthy agents was determined. Our results show that alternative credit scoring for these agents is possible. To our knowledge, no paper has investigated how to determine the creditworthiness of the population used in this project (agents who help customers pay for goods and services).

References

[1] S. M. Sadatrasoul, G. R. Mohammad, M. Siami, and Z. Hajimohammadi, "Credit scoring in banks and financial institutions via data mining techniques: A literature review.," *Journal of AI and Data Mining.*, vol. 1, pp. 119–129, 2013.

[2] M. Hurley and J. Adebayo, "Credit scoring in the era of Big Data," *Yale Journal of Law and Technology*, vol. 18, no. 1, 2017.

[3] "What lenders really think about unscoreables," *THE SCORE*. [Online]. Available: https://thescore.vantagescore.com/article/67.

[4] "What's in my FICO® Scores?," *myFICO*. [Online]. Available: https://www.myfico.com/credit-education/whats-in-your-credit-score.

[5] J. Kagan, "VantageScore," *Investopedia*, 29-Jan-2020. [Online]. Available: https://www.investopedia.com/terms/v/vantagescore.asp.

[6] "Global Findex," *Global Findex*. [Online]. Available: https://globalfindex.worldbank.org/#about_focus.

[7] "What is mobile money? All about Mobile Money guide from WorldRemit," *WorldRemit*. [Online]. Available: https://www.worldremit.com/en/mobile-money.

[8] D. Musiime and S. Schueth, "TANZANIA QUICKSIGHTS REPORT FIFTH ANNUAL FII TRACKER SURVEY," *Financial Inclusion Insights*, Jun-2018.

[9] R. Aitken, " 'All data is credit data': Constituting the unbanked," *Competition & Change*, vol. 21, no. 4, pp. 274–300, 2017.

[10] B. Baesens, T. V. Gestel, S. Viaene, M. Stepanova, J. Suykens, and J. Vanthienen, "Benchmarking state-of-the-art classification algorithms for credit scoring," *Journal of the Operational Research Society*, vol. 54, no. 6, pp. 627–635, 2003.

[11] S. Lessmann, B. Baesens, H.-V. Seow, and L. C. Thomas, "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research," *European Journal of Operational Research*, vol. 247, no. 1, pp. 124–136, 2015.

[12] N. Hsieh, "Hybrid mining approach in the design of credit scoring models," *Expert Systems with Applications*, vol. 28, no. 4, pp. 655–665, 2005.

[13] N.-C. Hsieh and L.-P. Hung, "A data driven ensemble classifier for credit scoring analysis," *Expert Systems with Applications*, vol. 37, no. 1, pp. 534–545, 2010.

[14] Q. Yu, Y. Luo, C. Chen, and X. Ding, "Outlier-eliminated k-means clustering algorithm based on differential privacy preservation," *Applied Intelligence*, vol. 45, no. 4, pp. 1179–1191, Nov. 2016.

[15] A. Blanco, R. Pino-Mejías, J. Lara, and S. Rayo, "Credit scoring models for the microfinance industry using neural networks: Evidence from Peru," *Expert Systems with Applications*, vol. 40, no. 1, pp. 356–364, 2013.

[16] A. Shema, "Effective credit scoring using limited mobile phone data," Proceedings of the Tenth International Conference on Information and Communication Technologies and Development - ICTDX 19, 2019.

[17] D. Shen and Z. Lu, "Computation of Correlation Coefficient and Its Confidence Interval in SAS®," *SUGI 31*, pp. 170–31.

[18] C. H. Li and S. C. Park, "An efficient document classification model using an improved back propagation neural network and singular value decomposition," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3208–3215, 2009.

[19] G. W. Furnas, S. Deerwester, S. T. Durnais, T. K. Landauer, R. A. Harshman, L. A. Streeter, and K. E. Lochbaum, "Information Retrieval using a Singular Value Decomposition Model of Latent Semantic Structure," *ACM SIGIR Forum*, vol. 51, no. 2, pp. 90–105, Feb. 2017.

[20] G. Hackeling, Mastering machine learning with scikit-learn: learning to implement and evaluate machine learning solutions with scikit-learn. Birmingham, UK: Packt Publishing Ltd., 2017.

[21] M. Capó, A. Pérez, and J. A. Lozano, "An efficient approximation to the K-means clustering for massive data," *Knowledge-Based Systems*, vol. 117, pp. 56–69, 2017.

[22] W. L. Hamilton, "Applied Machine Learning, Lecture 18: Unsupervised Learning," *https://cs.mcgill.ca/~wlh/comp551/*, 2019.

[23] "sklearn.cluster.KMeans," *scikit learn*. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html.

[24] F. Louzada, A. Ara, and G. B. Fernandes, "Classification methods applied to credit scoring: Systematic review and overall comparison," *Surveys in Operations Research and Management Science*, vol. 21, no. 2, pp. 117–134, 2016.

[25] "What is Transaction Data and Who Uses It?," *INETCO*. [Online]. Available: https://www.inetco.com/monitor-transactions/.

[26] A. Mosavi, P. Ghamisi, Y. Faghan, and P. Duan, "ComprehensiveReviewof Deep Reinforcement Learning Methods and Applications in Economics," *Mathematics*, vol. 8, no. 10, p. 1640, Sep. 2020 [Online]. Available: http://dx.doi.org/10.3390/math8101640

[27] V. Ha. and H. Nguyen, H., "Credit scoring with a feature selection approach based deep learning," *MATEC Web of Conferences*, vol 54, p.05004, Apr. 2016 [Online]. Available: 10.1051/matecconf/20165405004

[28] P. Addo, D. Guegan, and B. Hassani, "Credit Risk Analysis Using Machine and Deep Learning Models," *Risks*, vol. 6, no. 2, p. 38, Apr. 2018 [Online]. Available: http://dx.doi.org/10.3390/risks6020038