An automated credit intelligence learning system

Tri Handhika*, Siti Fatimah and
Muhammad Farkhan Novianto

Centre for Computational Mathematics Studies,
Gunadarma University,
Depok, 16424, Indonesia
and
Data Science Department,
Amartha Mikro Fintek,
Jakarta, 12560, Indonesia
Email: trihandika@staff.gunadarma.ac.id
Email: siti.fatimah@amartha.com
Email: farkhan.novianto@amartha.com
*Corresponding author

Abstract: To accelerate the financial services, microfinance requires tools and technologies to provide an automated dynamic credit decision which leads to an accountable and efficient system. Considering a case on loan disbursement in the micro-business sector, this study presents a very comprehensive innovation, namely automated credit intelligence learning system (ACILES) which consists of dynamic credit scoring and optimal dynamic credit pricing: derived from tenor, rate, installment and plafond (TRIP). While credit pricing is obtained from the profit based pricing and simulation process, the credit scoring is developed by modelling not only the borrower's profile, but also psychometric analysis of the perception of borrower and surveyor via item response model which is combined with multivariate adaptive regression splines (MARS) model and structural equation modelling (SEM), respectively. By performing the experiment, it is clearly proved that ACILES can be implemented in order to augment microfinance business capacity.

Keywords: automated; credit pricing; credit scoring; dynamic; learning system.

Reference to this paper should be made as follows: Handhika, T., Fatimah, S. and Novianto, M.F. (2022) 'An automated credit intelligence learning system', *Int. J. Electronic Finance*, Vol. 11, No. 2, pp.87–100.

Biographical notes: Tri Handhika received his BSc and MSc in Mathematics and Natural Sciences, majoring in Mathematics especially in Actuarial Statistics and in Computational Finance from the University of Indonesia, Depok, Indonesia in 2008 and 2011, respectively, and PhD in Computer Science especially in High-Performance Computing from the Gunadarma University, Depok, Indonesia in 2015. Currently, he is a researcher at the Centre for Computational Mathematics Studies, Gunadarma University, Indonesia. He is also a senior data science consultant for several technology companies in Indonesia to build an automated recommendation system to support data-driven decision management. His research interests include stochastic modelling, actuarial mathematics, numerical simulation, data science and high-performance computing.

Siti Fatimah received his Bachelor's in Mathematics and Natural Sciences, majoring in Mathematics from the University of Indonesia, Depok, Indonesia in 2014. Currently, she works at the Amartha Mikro Fintek in Data Science Team. She has experience in business intelligence, statistical modelling, and customer analysis.

Muhammad Farkhan Novianto is an Experienced Data Scientist with a demonstrated history of working in the internet industry. He is skilled in statistical modelling, data science, market research, predictive analytics and Hadoop. He is a strong engineering professional with focused in his Bachelor's in Statistics from the Universitas Gadjah Mada (UGM).

This paper is a revised and expanded version of a paper entitled 'An automated credit intelligence learning system' presented at International Conference on Data Science, Yogyakarta, Indonesia, 3–5 December 2016.

1 Introduction

Microfinance serves as an avenue of providing financial access to low-income individuals or groups in order to encourage economic development in a given area. According to the World Economic Forum (2015), governments, state agencies and international organisations have made several attempts to improve the current economic situation by acknowledging the importance of small businesses. Therefore, with a group-based lending system that facilitates loan disbursements in micro-business (group liability), microfinance has successfully solved multiple financial constraints. Furthermore, microfinance allows others to become independent, which is imperative when managing one's own business.

While emphasising the need for financial inclusion, it is important to understand that traditional microfinance methodologies do not serve the best approach to accelerate process of loan disbursement to people in rural areas anymore. After measuring the impact of microfinance institutions, it is interesting to observe several obstacles such as accessibility, time, cost efficiency and area coverage. This financial disbursement challenge has triggered microfinance institutions to transform their movement, from conventional to technology-based methods of financial distribution.

The process of financial technology has omitted several constraints of conventional lending process with financial service enhancement and integrated connections throughout peers and institutions such as finding new customers and deepening customer relationships. It also provides access to build a store-of-value with peer-to-peer lending solution using data analytics to develop products. Therefore, the concept of financial technology for micro-business is becoming increasingly relevant.

Through credit scoring (CS) and credit pricing information such as tenor, rate, installment and plafond (TRIP) derived from peer-to-peer lending, we can implement a higher demand of access acceleration and immediate disbursement. This is due to the fast process of customising the borrower's risk calculation. The main differences between conventional and technology-based systems are in the usage of data. An effective online lending system via machine learning leverages a psychometric method during its CS analysis and risk assessment, to understand how to efficiently tackle the difficulties that exist among the borrowers. Additionally, online loan applications with CS-TRIP avoid that the challenges in-person services face by carrying out the process remotely.

Several recent researches related to the use of CS in improving credit disbursement performance, including research on CS by fusing social media information in online peer-to-peer lending as conducted by Zhang et al. (2016); stress tests to analyse the impact of economic changes for the effective management of a technology credit fund using behavioural technology CS model with time-dependent covariates as researched by Ju et al. (2015); and the importance of alternative CS used by Indonesia's Fintech Lending in driving economic growth through financial inclusion as reported by PwC Indonesia (2019).

This proposed system combines qualitative and quantitative method dynamically in order to configure the appropriate strategies. Investors can rely on the results of the CS-TRIP to help them choose more reliable borrowers to invest in and decrease their overall risk profile. Investors can also observe their business growth in real-time without any additional constraints. This proposed automated system is far more effective than one based purely on subjectivity and human judgement. In light of this fact, it becomes important to develop an automated and dynamic CS system to accelerate the loan disbursement process.

2 Related works

CS is a system for predicting borrower eligibility to earn loan disbursement, which is useful in determining whether an applicant is creditworthy or not. The measurement of this system is based on the likelihood of a loan defaulting on its financial obligation. CS has become one of the most crucial technologies that may affect performance management in handling microfinance business. It indicates a borrower's eligibility to obtain a loan or not based on their credit history. However, CS can also be used for informal sector clients who lack credit histories and have no scores on other indicators as mentioned in Rhyne and Christen (1999).

An experiment conducted by Schreiner (2000) in Bolivia showed that the configuration of CS in microfinance improved the judgement of credit risk. In addition, the operational cost of lending can also be significantly reduced. Hence, the implementation of credit system in microfinance has successfully resulted in several benefits, such as enhancement in efficiency, profitability and market share, reduction of cost as well as losses, and professional image management as stated by Vega et al. (2013).

CS plays an important role to quantify the credit risk factor. In order to evaluate the accuracy of this method, various prediction techniques have been formulated and introduced. Several measurement methods are discussed in Yu et al. (2015) such as expert systems, econometric models, artificial intelligence (AI) and hybrid form. The expert systems are used as a basic method to assess credit risk through subjective analysis, which is highly dependent on its subjective judgement. Equally, econometric models use the quantitative methods for data analysis and prediction based on mathematics, statistic, and computer science. Hereinafter, with its fundamental techniques such as artificial neural networks (ANNs) and support vector machines (SVM), AI becomes the most powerful computational learning tool to configure the human intelligence process. ANNs itself, has been widely used to measure credit score. Hybrid form is a combination of two or more methods for measuring the strengths and minimising the weaknesses of each method.

Many researchers have conducted an accurate comparison of several CS methods. Desai et al. (1996) used a multilayer perceptron neural networks (NNs) for CS and found that NN models outperform linear discriminant analysis (LDA) and logistic regression (LR) models when the performance measurement is percentage of bad loans correctly classified. Lee et al. (2006) found that the classification and regression tree (CART) and multivariate adaptive regression splines (MARS) outperform traditional discriminant analysis (DA), LR, NNs, and SVMs in terms of CS accuracy.

Another study by Lee et al. (2002) proposed an integration between CS model using back propagation neural networks (BNN) with a traditional DA approach and implied that the proposed hybrid approach converged much faster than the NNs model. The study also outperformed traditional DA and LR methods. On the other hand, Lee and Chen (2005) proposed a two-stage hybrid CS model employing ANN and MARS and found that this hybrid model successfully outperformed LDA, LR, single ANNs and single MARS. Nevertheless, research in terms of accuracy of CS models is still ongoing.

To solve the credit classification problems, several studies have focused on reducing the type-1 and type-2 errors from their models. These models switch the rejected good credit applicants and reassign them to a conditionally accepted class. In order to evaluate these models, Chen and Huang (2003) proposed a hybrid methodology by applying NNs and genetic algorithm (GA) and stated that the proposed hybrid model was a potentially effective tool to reassign the rejected applicants to the preferable accepted class, which used customer balance adjustments between costs and preferences.

On the other hand, Chuang and Lin (2009) presented a reassigning credit scoring model (RCSM) to solve classification problems and decrease type-1 error by developing an ANNs and case-based reasoning (CBR) based on CS model for reassignment of the rejected good credit applicants to the conditionally accepted class. In this study, the proposed hybrid model was more accurate compared to the other CS methods that were commonly used. Hence, this study also contributed to the reduction of type-1 error in the scoring system. Furthermore, a study by Li and Zhong (2012) used various techniques for CS in order to maximise revenue and reducing the type-1 and type-2 errors. The study compared several statistical models such as LDA, LR, MARS, Bayesian model, and decision tree and also various AI models such as ANN, GA, and SVM. In addition, other methods including hybrid method and ensemble method as in Handhika et al. (2019) were also applied in the study.

The rapid development of CS methods has the potential to be adopted in microfinance, as previously analysed by Schreiner (2000) when implementing CS in estimating risk of microfinance's borrower quantitatively, not qualitatively as is commonly applied. The main differences of microfinance when compared to formal financial institutions is in the flow of information, where microfinance usually used qualitative method and informal approach. In a further analysis, Schreiner (2004) suggests that the quantitative method is able to help microfinance because the risk prediction is better than the all-good 'Naïve' model currently used by the microfinance in which all loans approved by the traditional evaluation process are disbursed.

However, CS for microfinance is still less powerful when compared to the scoring for formal financial institution products. Thus, the knowledge of information by qualitative method is still needed. By using information from the qualitative methods, microfinance can predict and measure credit risk based on repayment behaviour and this repayment behaviour analysis can be done using psychometric tools.

Klinger et al. (2013) carried out a pilot test of an innovative psychometric tool. This pilot test aimed to evaluate credit risk of business owners in Peru seeking a loan to develop business. In this approach, the psychometric tool compared the behaviour of business owners with those in other countries that applied the same psychometric tool. Results of this comparison pointed out that, despite the personality differences of business owners among countries, the dimensions of business performance and their credit risk are common. The psychometric CS was discovered for the first time by Arraiz et al. (2015) and results showed that the psychometric test can lower the risk of the loan portfolio for bankable entrepreneur, which is derived from their credit history. This mechanism serves as the secondary screening platform for their credit. According to Entrepreneurial Finance Lab, the use of psychometric tests increased access to credit for unbanked entrepreneurs without any risk affecting the loan portfolio.

To maximise profitability, business owners need to understand the elasticity from customers' demand. Thus, the idea of profit based pricing is based on profit maximisation with a combination of cost, risk, and price elasticity of customers by referring to Experian Decision Analytics (2013). Profit based pricing in Phillips (2013) is used to determine optimal prices from any given pricing segment, using objective function containing the net interest income (NII).

As mentioned above, price based segment is derived from credit risk strategy; which did not provide any numerical result. Oliver and Oliver (2014) successfully found a numerical algorithm of price based segment. This algorithm is used to find optimal prices to maximise return on equity by considering price response and default risk, based on the nonlinear differential equations.

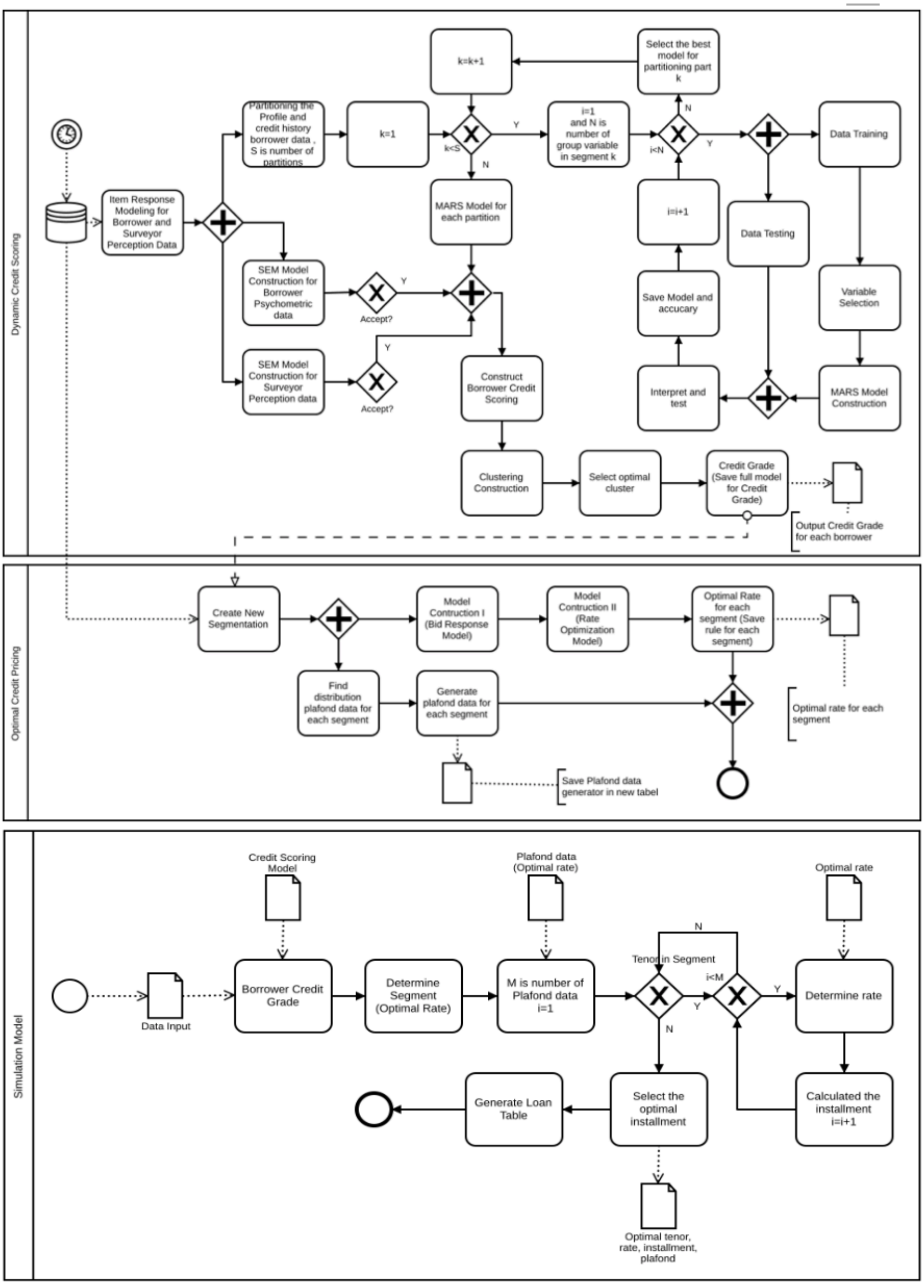
The main purpose of adopting an analytical approach to credit pricing is for increasing profitability. Boyd et al. (2005) reported that a UPS package shipper earned an increase in profit, which more than $100 million per year over previous business practices by using target pricing system (TPS). Another study by Phillips (2010) reported that the profitability of the sub-prime auto lender AmeriCredit is also increased about $4 million in three months by implementing this credit pricing optimisation system. TPS itself, is a bid pricing system that considers cost, price sensitivity, and competitive environment factors.

The three main elements of TPS function as a measurement variable when generating bids of goods portfolio or service that perform in a contract period. Logically, this is the reason that many companies utilise credit pricing to improve both the processes and the systems that they use to set the prices.

3 An automated credit intelligence learning system

Due to the challenge to accelerate loan disbursement on microfinance, an automated credit intelligence learning system (ACILES) is created to assist arduous conditions between lender and borrower on the process of lending. ACILES itself, is a machine that has the capacity to gather and analyse loan disbursement by micro-business through data aggregation from its loan portfolio. This system has also become a platform to integrate the communication process between the systems used. Based on the configuration of machine learning, these systems are also able to adapt with recent data within a second, which allow business owners to monitor business performance and also perform the managerial process remotely.

Figure 1 An ACILES



In its process, ACILES combines dynamic CS process to assess the creditworthiness of the applicant and its risks as well as the process of determining automatic credit pricing. A credit score is constructed based on qualitative and quantitative data, in order to improve the measurement of credit risk in microfinance business. Subsequently, a machine is constructed to determine the credit pricing for each risk. Therefore, ACILES aims to accelerate the process of credit analysis, so that the process of loan disbursement becomes faster and safer, which improves microfinance business sustainability.

ACILES process is divided into two stages as displayed in Figure 1. The first is automated model construction process. In this stage, the machine learning creates a dynamic CS and optimal credit pricing models which consist of four variables: TRIP. The CS model consists of two components: item response model, which combines MARS models based on the borrower's qualitative and quantitative data; structural equation modelling (SEM), which is aggregated by psychometric analysis of the perception of both borrower and surveyor. This machine will run automatically and periodically in order to accommodate data that grow rapidly. In the learning process, the machine will study the risks based on existing data and determine the credit price to maximise profitability, which corresponds to each risk. In this process, the score that was updated regularly will allow the same borrower to obtain different pricing and credit score for any given period.

The second stage is real-time CS-TRIP simulation process. Based on the model in the first stage, the output from this second stage process will deliver a prediction of the CS- TRIP and this prediction includes a degree of risk and pricing that will be given to the loan applicant. Subsections 3.1–3.3 describes each process from Figure 1 in more details.

3.1 Dynamic CS

CS models are initiated in order to measure creditworthiness aspect of borrowers. According to the dynamic CS data from Figure 1, the process is divided into three main stages.

The first stage involves creation of borrower portfolio using item response model that is combined with MARS model. MARS model is chosen to configure this system because of its flexibility to handle continuous and categorical data, compared with a linear regression model. In addition, linear relationship between dependent and independent variables are not required in MARS model. Furthermore, compared to AI methods, MARS is more superior because of its short training time and compatibility to interpret as discussed by Li and Zhong (2012). MARS is nonlinear and nonparametric regression method with high-dimensional data, as proposed by Friedman (1991).

In the second stage, psychometric data will be processed using SEM model to identify borrower's perception. The third stage is similar to the second. However, it involves identifying surveyor's perception.

SEM is used to describe the pattern of the relationship or correlation between the set of variables in the model. SEM is a collection of statistical techniques to examine relations between one or more independent and dependent variable. Both independent and dependent variables can be used to measure unobserved latent variable, which is not directly observed. This form of SEM is divided in two forms: continuous and discrete as stated by Ullman (2006). In its fundamental base, SEM can be viewed as combination of factor analysis and linear regression or path analysis. These relationships between latent variables are represented by regression. Furthermore, relationship between observed variable can be examined by covariance structure modelling.

Borrower data consists of variables, where data redundancy is likely to happen. If the data is directly used in the regression model, the multicollinearity of emergence will be allowed in this process. The variables are grouped in such a way that there are no redundant variables. This condition facilitates to reduce multicollinearity and such a category is referred to as a combination of variables. Moreover, this combination is used for the initial variable selection in the modelling process in order to obtain the best models that are parsimonious. Rismayanti (2009) divided credit into two main types; productive and consumptive credit. Productive credit are the loans that are used to finance the working capital needs of borrowers such as to facilitate the production, whether in agriculture, industry, trade and other productive sectors. Conversely, consumer credits are the loans that are granted to fulfill the consumptive needs of borrowers, such as financing of education and house renovation. The credit objectives of the borrowers will affect credit ratings during the credit analysis process. Therefore, a different model is imperative to fulfill different purposes.

The process of modelling using MARS method begins by dividing the data based on credit objectives. This data will be modelled from a combination of variables that have been formed in each part. Therefore, the optimal MARS model is selected in a two-stage process. In the first place, a very large number of basic functions are applied to fit the initial data. In the second stage, the basic functions are deleted based on their level of contribution. This backward stepwise process is using criteria from generalised cross validation (GCV). From GCV, the best model (the greatest R2) of variable combination will be chosen. All in all, it can clearly be seen that each credit objective has its different MARS model.

At the same time, SEM models the psychometric data separately based on the perception of both borrower and surveyor. The process of modelling is summarised in four stages: model specification, model estimation, model evaluation, and model modification. Firstly, model specification consists of stating the hypotheses to be tested in both diagram and equation form. Statistical identification and evaluation is also used to evaluate the underlying assumption regarding the model. Secondly, the relation in the diagram is also translated into equations in order to estimate the model. To illustrate, one of the method in model specification is in the Bentler and Weeks (1980) method. Thirdly, model evaluation is done to assess a suitable model in SEM and to interpret parameter estimation. The last stage is model modification. This model is used to test the hypothesis in a theoretical work and to improve the analysis transitions from confirmatory to exploratory. If the parameter in regression part of SEM is positive, the result of SEM model is compatible to be used to improve credit score predictions. Furthermore, the results of MARS and SEM model are combined with a certain measurement scoring to build a decent credit score where hierarchical clustering (HC) algorithm is used to obtain the credit grade. This algorithm was chosen because it has several advantages including embedded flexibility regarding a level of granularity, ease of handling of any forms of similarity or distance, and it is more versatile by referring to Abbas (2008).

3.2 Optimal credit pricing

Differences in risk profiles of borrowers is the reason that many lenders do not offer the same price of credit. Phillips (2013) stated that there are two reasons for charging different prices to its different segments. Firstly, incremental cost and risk vary in each segment. Most of the lenders differentiate prices, according to product and borrower characteristics as well as channel. Secondly, borrower price sensitivity is different for each segment. This study will try to separate the borrower based on their amount of plafond, tenor, CS, and the k-funding.

The main challenge in price optimisation is to determine the exact rate of annual percentage rate (APR) to be charged for each segment. This price should be updated regularly in order to respond to market changes, cost of funds or competitive action. Therefore, price optimisation involves a probability for loan funding with incremental profitability to lenders. An increase rate for potential borrower reduces the likelihood that the customer receives a loan, but it increases profits when the customer accepts it. This loan profitability is measured by NII. On the other hand, present value net interest income ($PVNII$) measures the expected present value of the loan after tax payment or benefit entirely from lending.

In the funding process, there are possibilities that a borrower defaults, which causes the payment obstructed. This condition mostly happens during the loan period. Assume that the lender pays a periodic rate r_c for capital and has an internal discount rate r_d. Let s_i be the probability that a borrower makes a payment to period i. Also let P as the amount borrowed, r and n is the rate and the term, respectively. Then the expected value of NII is:

$$\begin{aligned} &PVNII(P, r, n) \\ &= P\left[\Sigma_{i=1}^{n}\left(\frac{1}{(1+r_d)^i}\right)\left(\frac{s_i r(1+r)^n}{(1+r)^n-1}-\frac{r_c\,(1+r_c)^n}{(1+r_c)^n-1}\right)\right] \end{aligned} \tag{1}$$

where $s_i = \Pi_{j=1}^{i}(1-p_j)$ for $i = 1, 2, ..., n$ and p_j is the probability that the borrower defaults in period j.

The following is the price optimisation problem in Phillips (2013) that maximise expected total profitability:

$$\max_r TR(r) = \Sigma_i D_i \bar{F}_i(r_i)\left[PVNII(P_i, r_i, n_i) + v_i\right], r_i \geq 0 \tag{2}$$

where $N \geq 1$ is the number of pricing segment, $r = r_1, r_2, ..., r_N$ is the vector of rates offered to each pricing segment, $D_i > 0$ is the total demand (in number of loans) in pricing segment i, $P_i > 0$ is the average loan size in pricing segment i, $n_i \geq 1$ is the typical term in pricing segment i, $PVNII(P_i, r_i, n)$ is the present value of NII, v_i is the present value of expected non-interest items (such as fees and operating cost) in pricing segment i, $\bar{F}_i(r_i)$ is the bid responses model, the fraction of successful applicants who will accept the loan as a function of the rate r_i. $\bar{F}_i(r_i)$ is a function such as the probit or logit that can be interpreted as complementary cumulative distribution function of a probability distribution.

The optimisation problem can be solved using nonlinear programming. $PVNII(P_i, r_i, n)$ is log-concave in r and that is sufficient for existence and uniqueness of solution, which pointed out by Phillips (2013). Log-concavity is a weaker condition than concavity.

At this stage, the system is not only suitable to get the optimal rate, but also to generate plafond table that is used in the simulation process. The table is obtained by determining the most appropriate distribution of plafond for each segment. Each distribution of the segment generates a random number, which is quite large. In the end, a complete plafond table for each segment is obtained.

3.3 Simulation process

The simulation process is used to measure the creditworthiness of a borrower by determining the rate that will be given to new loan applications based on CS model and by the previous optimal credit pricing, i.e., TRIP. This process consists of five stages as displayed in Figure 1.

The process starts with the determination of a borrower's credit grade based on CS model and loan application data. At the second stage, the determination of a borrower segment based on the loan application and borrower's credit grade. At the third stage, each plafond in the plafond table is determined in accordance with the rate of borrower segment. At the fourth stage, installment is calculated based on the given rate of borrower segment for each plafond. At the final stage, the optimal installment is selected based on the installment closest one with borrower's willingness to pay.

Due to the obstacles to get a valid data of the borrower, the amount that is given to the borrower is based on borrower's willingness to pay. At this stage, we also obtain the optimal TRIP in order to create a measurement of borrower's ability to make a repayment. Therefore, the loan table contains a selection of the plafond and tenor, in order to see which plafond value is close to optimal plafond. The table also displays installment to be paid according to the selected loans.

4 Results and discussion

This study uses panel data from 2014–2020 via PT. Amartha Mikro Fintek to implement ACILES for calculating CS and credit pricing such as TRIP. Data is divided into several components: borrower profile, loan disbursement data, and credit dues. The study also uses psychometric data on the perception of both borrower and surveyor, in order to improve the information about borrower. Therefore, the items of data used are based on the result by item response modelling. Item response models are to enhance conceptual toolkit and research technique by its researchers. This method is meant to gain a better level of understanding in psychometric measurement. Furthermore, data is analysed using the same process described in Figure 1, which runs automatically every week. The output from this credit score model is a form of mathematical equation, which is used to predict credit score from new borrowers.

Meanwhile, the other output of this credit pricing is optimal rate value for any given segment shown in Table 1. Based on data collected over the last seven years, there are 84 segments formed starting from Rp.3,000,000 until Rp.15,000,000 with a range of Rp.500,000 per segment. The tenors imposed range from 4, 10, 25 and 50 weeks, with several variations of credit grade and the number of times the funding has been received by the borrower. The optimal weekly rate is obtained from the calculation of equation (2) using information from each segment.

Table 1 Optimal rate per segment

No.	*Plafond*	*Tenor (weeks)*	*Credit grade*	*The k-funding*	*Weekly rate*
1	Rp.14,500,000–Rp.15,000,000	50	A	2	0.58%
2	Rp.14,000,000–Rp.14,500,000	50	A-	3	0.53%
3	Rp.13,500,000–Rp.14,000,000	50	B	2	0.66%
4	Rp.13,000,000–Rp.13,500,000	50	B	2	0.64%
5	Rp.12,500,000–Rp.13,000,000	50	C	2	0.72%
...	...	...	...	...	...
84	Rp.3,000,000–Rp.3,500,000	50	A	1	0.52%

Figure 2 Loan table of Hartini (simulation model) (see online version for colours)


Name : Hartini
Credit Grade : A-

Plafond	Tenor (in week) 4	10	25	50
Rp.4.200.000	Rp.901.000 (2.81%)	Rp.379.000 (6.23%)	Rp.187.000 (14.83%)	Rp.110.000 (30.92%)
Rp.4.300.000	Rp.927.000 (2.92%)	Rp.389.000 (6.04%)	Rp.192.000 (15.07%)	Rp.112.000 (30.14%)
Rp.4.400.000	Rp.952.000 (2.96%)	Rp.400.000 (6.15%)	Rp.197.000 (14.65%)	Rp.115.000 (30.68%)
Rp.4.500.000	Rp.978.000 (2.98%)	Rp.410.000 (6.31%)	Rp.202.000 (14.91%)	Rp.118.000 (31.28%)
Rp.4.600.000	Rp.1.004.000 (2.92%)	Rp.420.000 (6.12%)	Rp.207.000 (15.23%)	Rp.120.000* (30.32%)
Rp.4.700.000	Rp.1.030.000 (2.97%)	Rp.430.000 (6.19%)	Rp.212.000 (14.79%)	Rp.123.000 (30.86%)

Figure 3 Loan table of Hartono (simulation model) (see online version for colours)


Name : Hartono
Credit Grade : B

Plafond	Tenor (in week) 4	10	25	50
Rp.4.200.000	Rp.1.112.000 (5.90%)	Rp.463.000 (10.24%)	Rp.195.000* (16.07%)	Rp.123.000 (46.43%)
Rp.4.300.000	Rp.1.137.000 (5.77%)	Rp.470.000 (9.30%)	Rp.200.000 (16.27%)	Rp.126.000 (46.51%)
Rp.4.400.000	Rp.1.163.000 (5.73%)	Rp.478.000 (8.63%)	Rp.205.000 (16.48%)	Rp.128.000 (45.45%)
Rp.4.500.000	Rp.1.189.000 (5.69%)	Rp.485.000 (7.78%)	Rp.212.000 (17.78%)	Rp.131.000 (45.55%)
Rp.4.600.000	Rp.1.216.000 (5.74%)	Rp.496.000 (7.83%)	Rp.219.000 (19.02%)	Rp.132.000 (43.48%)
Rp.4.700.000	Rp.1.241.000 (5.62%)	Rp.506.000 (7.66%)	Rp.226.000 (20.21%)	Rp.135.000 (43.62%)

On the other hand, the simulation process shown in Figure 1 occurs in all new transactions. For a deep dive, considering Hartini proposed a loan amounting Rp.4,500,000. For repayment, Hartini commits to pay Rp.125,000 every period. Based on loan filing provided, the ACILES will directly count Hartini's credit score and its optimal credit pricing. Figure 2 shows the loans that can be taken by Hartini as the output of the ACILES. In this case, the plafond recommended by system (*) is Rp.4,400,000, which will be repaid within 50 periods (weeks) with a repayment of Rp.115,000 at a total rate of 30.68%.

The system is set to recommend a loan, with the result that the loan repayment will be paid close to the value of the borrower's ability to pay. While loans will produce a maximum plafond, borrower can choose the loans beyond the recommendation. All rates generated in Figure 2 are the optimal rates. However, the main obstacle of this system is the difficulty to get a valid data for measuring a borrower's ability to earn credit automatically. Lastly, inputs from borrower are still required to validate the data.

The learning process of ACILES can be demonstrated by simulating the new application of a borrower (e.g., named Hartono) with exactly the same profile and psychometric data as Hartini, but for some time period ahead, e.g., two months later. Figure 3 shows different outcomes from those obtained by Hartini in Figure 2. This shows that the learning process of ACILES has been running properly every week. In addition, the difference shows that there has been a decreasing in the performance of installment payments by borrowers who have similar characteristics to Hartini during the last two months so that it affects the credit score and credit price of subsequent borrowers where previously Hartini got credit grade A – while Hartono got B.

5 Conclusions and future works

This study presents a project of an ACILES to accelerate the financial access in microfinance. The study develops a machine learning for dynamic CS, which functions to measure the creditworthiness of a borrowers' riskiness, and the process of determining the overall credit pricing. While the credit pricing itself considers TRIP. The study constructs a scoring model using an item response model that is combined with MARS model and SEM to obtain a more detailed information on the borrower. A decent credit score is clustered using HC to obtain a credit grade. Furthermore, a profit-based pricing model is used to obtain optimal credit pricing and the results show that the CS-TRIP can be implemented to accelerate operational process in one of the microfinance businesses in Indonesia, i.e., Amartha Mikro Fintek. For future research, the maintenance of ACILES becomes necessary and this involves validation to the current data, model evaluation by CS-TRIP, and other aspects such as including model improvement and model alteration with better accuracy. This idea is proposed in order to maintain and accelerate ACILES performance which will allow the system to analyse the borrower's credit efficiently. Most microfinance businesses use group liability system to secure the borrowing process and gain the loan guarantees as implemented by Grameen Bank and also by Amartha Mikro Fintek. Furthermore, in order to minimise the risk of greater losses, a predictive model of partial liability becomes more and more crucial and this model can be used to maximise group rate transfer as proposed by Allen (2016).

Thus, the model is suitable for predicting default rate from group lending and also for estimating extra funds needed to cover a borrower's default which is known as a group reserve valuation.

Acknowledgements

This document is the results of the research project funded by Direktorat Riset dan Pengabdian Masyarakat Deputi Bidang Penguatan Riset dan Pengembangan Kementerian Riset dan Teknologi/Badan Riset dan Inovasi Nasional (Grant No. 234/SP2H/LT/DRPM/2021).

References

Abbas, O.A. (2008) 'Comparisons between data clustering algorithms', *The International Arab Journal of Information Technology*, Vol. 5, No. 3, pp.320–325.

Allen, T. (2016) 'Optimal (partial) group liability in microfinance lending', *Journal of Development Economic*, Vol. 121, pp.201–216.

Arraiz, I., Bruhn, M. and Stucchi, R. (2015) *Psychometrics as a Tool to Improve Screening and Access to Credit*, IDB Working Paper Series.

Bentler, P.M. and Weeks, D.G. (1980) 'Linear structural equation with latent variables', *Psychometrika*, Vol. 45, No. 3, pp.289–308.

Boyd, D., Gordon, M., Andersson, J., Tai, C.H.C., Yang, F., Kolamala, A., Cook, G., Guardino, T., Purang, M., Krishnamurthy, P., Cooke, M., Nandiwada, R., Monteiro, B. and Haas, S. (2005) *Target Pricing System*, United States Patent, No. US 6,963,854 B1.

Chen, M.C. and Huang, S.H. (2003) 'Credit scoring and rejected instances reassigning through evolutionary computation techniques', *Expert Systems With Applications*, Vol. 24, No. 4, pp.433–441.

Chuang, C.L. and Lin, R.H. (2009) 'Constructing a reassigning credit scoring model', *Expert Systems With Applications*, Vol. 36, No. 2, pp.1685–1694.

Desai, V.S., Crook, J.N. and Overstreet Jr., G.A. (1996) 'A comparison of neural networks and linear scoring models in the credit union environment', *European Journal of Operational Research*, Vol. 95, No. 1, pp.24–37.

Experian Decision Analytics (2013) *Price Optimization in Retail Consumer Lending*, An Experian Decision Analytics White Paper.

Friedman, J.H. (1991) 'Multivariate adaptive regression splines', *The Annals of Statistic*, Vol. 19, No. 1, pp.1–67.

Handhika, T., Fahrurozi, A., Zen, R.I.M., Lestari, D.P., Sari, I. and Murni (2019) 'Modified average of the base-level models in the hill-climbing bagged ensemble selection algorithm for credit scoring', *Procedia Computer Science*, Vol. 157, pp.229–237.

Ju, Y., Jeon, S.Y. and Sohn, S.Y. (2015) 'Behavioral technology credit scoring model with time-dependent covariates for stress test', *European Journal of Operational Research*, Vol. 242, pp.910–919.

Klinger, B., Khwaja, A.I. and LaMonte, J. (2013) *Improving Credit Risk Analysis with Psychometrics in Peru*, Inter-American Development Bank.

Lee, T.S. and Chen, I.F. (2005) 'A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines', *Expert Systems With Applications*, Vol. 28, No. 4, pp.743–752.

Lee, T.S., Chiu, C.C., Lu, C.J. and Chen, I.F. (2002) 'Credit scoring using the hybrid neural discriminant technique', *Expert Systems With Applications*, Vol. 23, No. 3, pp.245–254.

Lee, T.S., Chiu, C.C., Chou, Y.C. and Lu, C.J. (2006) 'Mining the customer credit using classification and regression tree and multivariate adaptive regression splines', *Computational Statistics & Data Analysis*, Vol. 50, No. 4, pp.1113–1130.

Li, X.L. and Zhong, Y. (2012) 'An overview of personal credit scoring: technique and future work', *International Journal of Intelligence Science*, Vol. 2, No. 4A, pp.181–189.

Oliver, B.V. and Oliver, R.M. (2014) 'Optimal ROE loan pricing with or without adverse selection', *Journal of the Operational Research Society*, Vol. 65, No. 3, pp.435–442.

Phillips, R. (2010) *Customized Pricing*, Columbia University Center for Pricing and Revenue Management Working Paper, No. 2010-1.

Phillips, R. (2013) *Optimizing Prices for Customer Credit*, Columbia University Center for Pricing and Revenue Management Working Paper, No. 2013-1.

PwC Indonesia (2019) *Indonesia's Fintech Lending: Driving Economic Growth Through Financial Inclusion*, Fintech Series.

Rhyne, E. and Christen, R. (1999) *Microfinance Enters the Marketplace*, USAID Microenterprise Publications.

Rismayanti, D. (2009) *Analisis Portofolio Kredit (Konsumtif dan Produktif) dan Pengaruhnya Terhadap Laba (Studi Kasus PT Bank X Tbk)*, Departemen Manajemen, Fakultas Ekonomi dan Manajemen, Institut Pertanian Bogor.

Schreiner, M. (2000) 'Credit scoring for microfinance: can It work?', *Journal of Microfinance*, Vol. 2, No. 2, pp.105–118.

Schreiner, M. (2004) 'Scoring arrears at a microlender in Bolivia', *Journal of Microfinance*, Vol. 6, No. 2, pp.65–88.

Ullman, J.B. (2006) 'Structural equation modeling: reviewing the basics and moving forward', *Journal of Personality Assessment*, Vol. 87, No. 1, pp.35–50.

Vega, M.D.C.D.L., Oliver, A.B., Mejias, R.P. and Rubio, J.L. (2013) 'Improving the management of microfinance institution by using credit scoring models based on statistical learning techniques', *Expert Systems With Applications*, Vol. 40, No. 17, pp.6910–6917.

World Economic Forum Report (2015) *The Future of FinTech: A Paradigm Shift in Small Business Finance*, World Economic Forum.

Yu, L., Li, X., Tang, L., Zhang, Z. and Kou, G. (2015) 'Social credit: a comprehensive literature review', *Financial Innovation*, Vol. 1, No. 6, pp.1–18.

Zhang, Y., Jia, H., Diao, Y., Hai, M. and Li, H. (2016) 'Research on credit scoring by fusing social media information in online peer-to-peer lending', *Procedia Computer Science*, Vol. 91, pp.168–174.