Journal of Sustainability Science and Management
Volume 17 Number 5, May 2022: 60-76
eISSN: 2672-7226

A NEW EFFICIENT CREDIT SCORING MODEL FOR PERSONAL LOAN USING DATA MINING TECHNIQUE FOR SUSTAINABILITY MANAGEMENT

RABIHAH MD. SUM[1], WAIDAH ISMAIL[1,3]*, ZUL HILMI ABDULLAH[1,2], NURUL FATHIHIN MOHD NOOR SHAH[1] AND RIMULJO HENDRADI[3]

[1]*Faculty of Science and Technology, Universiti Sains Islam Malaysia (USIM), 71800 Nilai, Negeri Sembilan, Malaysia.*
[2]*Faculty of Data Science & Information Technology, INTI International University, Nilai, Negeri Sembilan, Malaysia.*
[3]*Information System Study Program, Faculty of Science and Technology, Universitas of Airlangga, Indonesia Campus C, Surabaya, Indonesia.*

**Corresponding author: waidah@usim.edu.my*
Submitted final draft:15 October 2021 *Accepted: 11 March 2022* http://doi.org/10.46754/jssm.2022.05.005

Abstract: Credit scoring models are used in decision-making processes to produce an accurate prediction of an applicant's creditworthiness. A five-step credit scoring model for personal loans was developed using the seven-step credit scoring model by Siddiqi. It uses real data provided by a bank. This study aims to remove the unnecessary complexity of the credit scoring process. The five-step credit scoring model consists of data massaging, factor analysis, data mining modelling, credit scoring and post-modelling. To ensure accuracy, factors that were significant in determining the creditworthiness of applicants were used in the model, which are the type of installment, age, monthly expenses, job sector, payment method and income-to-finance ratio. Furthermore, by presenting a systematic and structured step for developing a credit scoring model, this study contributed to the research on credit scoring. Based on the findings of this study, banks may use this model to create their own credit scoring model to assess the creditworthiness of personal loan applicants. By managing risks with this model, banks can create a long-term solution for credit system management and aid in the decision-making process.

Keywords: Credit scoring, data mining, personal loan.

Introduction

A credit scoring model to assist loan officers in evaluating loan applications was developed in this study. It focuses on modelling the credit scores for applicants of personal loans and aims to assist banks in deciding to whom the bank should lend its money and key variables or factors contributing to the delinquent tendencies of a client's account.

Credit scoring is the assessment of risk associated with lending to an organisation or a consumer (an individual) (Edelman & Crook, 2002). According to Siddiqi (2012), credit risk scoring is a predictive tool. The tool can be used to evaluate the level of risk associated with loan applicants. Specifically, risk scoring tools calculate a loan applicant's risk score. The applicant's score is then used as the basis for decision-making on the loan application. In general, financial institutions in Malaysia use credit scoring to determine borrowers' creditworthiness to provide loans. The loans may in the form of personal, housing or car loans. Credit scoring is concerned with developing empirical models to aid in the decision-making process in the credit business (Crook *et al.,* 2007). A credit scoring model can also predict borrowers' probability of undesirable behaviour in the future (Baesens *et al.,* 2003).

This study presents the detailed process of the credit scoring modelling, including the systematic and structured steps in developing the model. The foundations of the development were based on the intelligent credit scorecard model developed by Siddiqi (2012). This study applies the techniques onto the real environment using personal loan applicants' data supplied by a Malaysian financial institution. Concerning this study, data mining was used to develop the credit scoring process for personal loans.

In addition, the data mining techniques were aligned with the statistical results. This main contribution of the study is the application of an established credit scoring model (Siddiqi, 2012) to real personal loan applicants' data provided by a bank. This study identified and uncovered a new credit scoring model from a transformed bank. Data used in this study was supplied by a bank that was initially a credit cooperative. Furthermore, this study aims to remove the unnecessary complexity of the credit scoring process. This study asserts that complex problems, such as evaluating an individual credit risk, do not necessarily need a complex solution; rather, they need to be disentangled and simplified to find the solution (Sum, R. M., 2015). Hence, a five-step credit scoring model was developed. The five steps are sufficient to enable loan officers to evaluate the credit risks of personal loan applicants.

Consumer Credit Scoring

Nowadays, banks, as financial institutions, thoroughly analyse customers' credit risks before allocating any financial resource. The use of credit scoring models can significantly improve customer selection and reduce credit risks (Ziemba *et al.,* 2020). Customers expect banks to designed financial products and services that suit them so that they can improve their social well-being and expand their economic activities (Salem & Mahamad, 2020). Hence, banks need to be active and manage their impact on social and environmental sustainability (Pentzlin, 2011). Sustainability thus represents an opportunity for banks to improve trust in the banking system. As for increasing trust on credit report scores, banks can have an active role in promoting financial education and improving financial awareness.

Arya *et al.* (2013) defined credit scores as a number that represents an assessment of the creditworthiness of a person, or the likelihood that the person will repay their debts. Credit scores are generated based on the statistical analysis of a person's credit report. The repayment of debt is contingent upon two factors: The ability to pay the debt and the borrower's willingness to pay. Arya *et al.* (2013) investigated whether personality factors contributed to a credit score based on a dependent variable — the Fair Isaac Cooperation (FICO) score. The FICO score is an internet-based credit score estimator (www.myfico.com) that provides a credit score range based on self-reported financial information.

Donga *et al.* (2012) stated that credit scoring was invariably used to answer one key question — the probability of default within a fixed period, usually 12 months. Depending on the information used during the modelling construction, credit scoring can be divided into application scoring and behaviour scoring. Logistic regression, linear regression, linear programming and classification trees are used in the banking industry to develop credit scorecard systems.

On the other hand, consumer credit scoring can be defined as a classification task, where clients receive either a good or bad credit status (Kruppa *et al.,* 2013). Hui *et al.* (2013) explained that personal credit scoring can be divided into three categories: Credit bureau scoring, application scoring and behaviour scoring. Credit bureau scoring is conducted by external credit scoring companies. There are three steps of the credit scoring process, which are processes before, processes at and processes after credit take-up (Kamleitner & Kirchler, 2007).

Processes before credit take-up refer to situations when consumers perform explicit information search, option comparison and deliberations before the decisions to take up credit or using their own funds are made. Miller (2015) investigated the impact of lenders' information on loan outcomes. The study found that borrowers' information was crucial in credit markets that catered sub-prime borrowers. Lenders' information was used to limit costly default by improving lender filtering. According to Miller (2015), information is a crucial component for credit markets to function well, as active screenings of loan applicants make it possible to redirect funds towards borrowers who are likely to repay them. This may reduce the private and social costs of loan defaults. Similarly, Ghosh (2018) studied the role of credit reporting systems in influencing bank loan delinquency by evaluating the effect of credit reporting system reform on non-performing loans. According to Ghosh (2018), better credit risk assessment helped mitigate some of the informational asymmetries involved in credit extension. Following this, addressing bad loan problems will ease the flow of credit. Higher non-performing loans can also dampen profitability, tie up bank capital and raise funding costs. A crucial challenge for policymakers is to address non-performing loans to unlock the credit supply and, ultimately, promote economic growth.

Credit Scoring Techniques

Since the 1930s, researchers on financial institutions have used statistical techniques as methods in their studies (Edelman & Crook, 2002). One of the areas that use statistical techniques is credit scoring (Abdou & Pointon, 2011).

Artificial intelligence is widely used for credit scoring applications. Some of the AI techniques used are the ensemble method (EM), artificial neural networks (ANNs) and supportvector machines (SVMs) (Sadatrasoul *et al.,* 2013). The EM is made up of a unique classifier with different parameters tuned or different combined classifiers. This technique has different types, such as bagging, boosting and random forests. In the case of credit scoring, different classifiers group an applicant and vote on the application for a final decision (Tan *et al.,* 2006). Banks use EM as a predictive technique to predict the credit risks of loan applicants (Chopra & Bhilare, 2018). ANNs, meanwhile, imitate human brains functionality and are commonly used in classification, clustering and optimisation problems (Tan *et al.,* 2006). ANNs recognise complex and non-linear patterns between input and output variables in credit scoring. These patterns are used by banks to predict the creditworthiness of new loan applicants. Specifically, banks use ANNs to determine the probability of loan applicants defaulting (Bequé & Lessmann, 2017). On the other hand, SVMs are based on statistical learning designed for binary classification that maximises the margins of separation between negative and positive datasets (Vapnik, 2000). Banks use SVMs as tools to create credit scoring models that also allow for non-linear relationships among variables influencing credit scoring (Ala'raj *et al.,* 2018).

Hsieh (2004a, 2004b) proposed a self-organisational neural network map to identify groups of customers based on repayment behaviour and frequency — monetary behavioural scoring predictors. The study combined two data sets — an effective credit card account information and the individual transaction records of these accounts. The study demonstrated that using a behavioural scoring model was useful for identifying customer characteristics and facilitating the development of marketing strategies. Furthermore, Sohn and Kim (2013) proposed a technology behavioural credit scoring model that reflected changes in borrowers' financial ratios. According to the study, financial institutions should collect the ratios after lending and use the ratios for behavioural credit scoring technology.

Apart from that, logistic regression was used to produce a credit scoring model based on four factors, management, technology, marketability and profitability (Sohn *et al.,* 2007). Nonetheless, the study found that strong correlations between individual evaluations attribute caused inaccurate predictions. The results also demonstrated that the accepted model had a sample selection bias. Kim and Sohn (2007) improved the performance of the logistic credit scoring model by using data from rejected and accepted loans.

Siddiqi (2012) developed a seven-stage credit scorecard process. The first stage is preliminaries and planning. At this stage, developers identify the purpose of the scorecard, team roles and responsibilities. The second stage is data review and parameter projection. At the second stage, developers determine relevant data, project parameters, loan performance window and loan performance category definitions. The third stage is development database creation. At this stage, developers develop sample specifications and adjust for prior probability or factoring. The fourth stage is scorecard development. At this stage, developers identify missing values and outliers, correlation and scorecard scaling. The fifth stage is scorecard management reports. The purpose of the report is to make operational decision, such as deciding the scorecard cutoff. The sixth stage is implementing the scorecard. However, before implementing the scorecard, developers need to conduct pre-implementation validation. The purpose is to test scoring accuracy and front-end evaluation. The results from the pre-implementation can also be used to set cutoff. The seventh and final stage is post-implementation. The two main areas at this stage are reporting and reviewing. Developers monitor and pinpoint and report the sources of delinquency and profit.

Research Methodology

This section demonstrates the steps taken in developing the credit scoring model. This study develops a five-step credit-scoring model for personal loans. The process was presented in Figure 1 and the explanation of the steps is as follows.

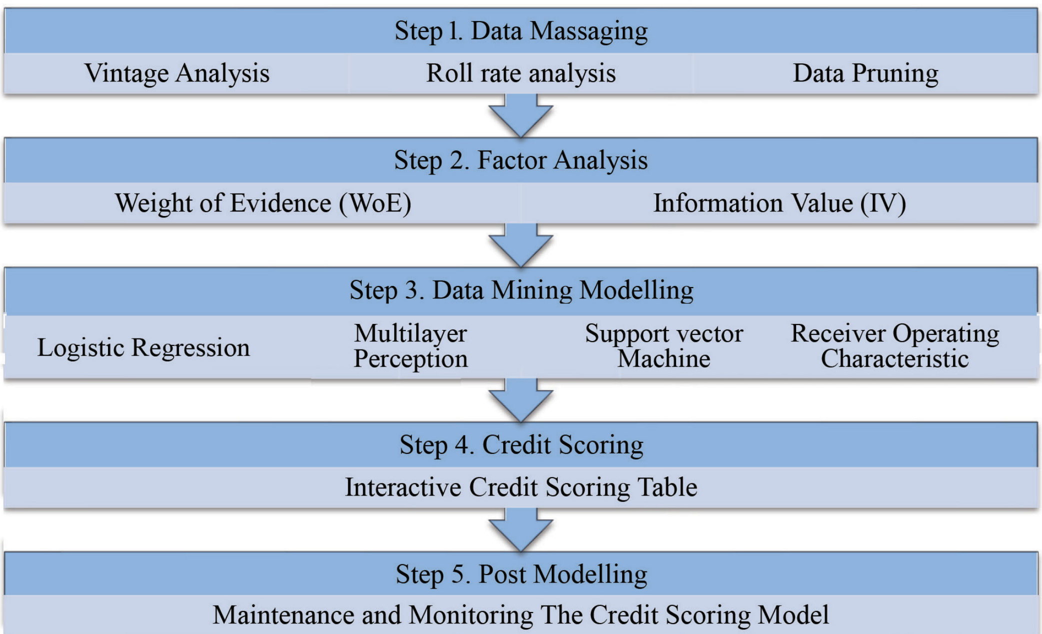


Figure 1: Steps in developing a credit scoring model for personal loans

Step 1: Data Massaging

Data massaging is the process of sorting and assessing raw data to ensure data usability and reliability. Data massaging is made up of three steps:

Step 1: *Vintage analysis* is conducted to evaluate the credit quality of a loan portfolio by analysing net charge-offs in the same originational period. This analysis is also used to determine the performance window.

Step 2: *Roll rate analysis* is conducted to determine the definition of a bad account based on the roll rates and the percentage of backward roll rates. Banks use roll rates to predict credit losses based on delinquency. Following this, analysing roll rates is an effective way to review overall trends and estimate future performance.

Step 3: *Data pruning* is where the data are filtered and omitted if they are not reliable. This is to ensure a successful analysis.

Step 2: Factor Analysis

Factor analysis comprises single- and multi-factor analyses. The single-factor analysis determines the strength of the relationship between individual key variables and the status of the account. The analysis consists of the computation of Weight of Evidence (WoE) and Information Value (IV). Meanwhile, the multi-factor analysis determines the interrelationship among key variables. The analysis is conducted using a correlation analysis.

$$\textit{Weight of Evidence (WoE)} = \left[ln \frac{Relative\ frequency\ of\ Goods\ (RFGs)}{Relative\ frequency\ of\ Bads\ (RFBs)}\right] * 100$$

$$WoE = 0, if \frac{RFGs = odds}{RFBs = 1}$$

$$WoE < 1, if\ RFBs > RFGs$$

$$WoE\ \geq 1, if\ RFGs\ \geq RFBs$$

$$\textit{Information Value (IV)} = \sum_{i-1}^{K}\left[(RFGs_i - RFBs_i) * ln\left(\frac{RFGs}{RFBs}\right)\right]$$

The IV of a predictors or a variable is the sum of the (absolute) values for WoE.

$$\textit{Correlation} = \frac{n(\sum xy)-(\sum x)(\sum y)}{\sqrt{[n\sum x^2-(\sum x)^2][n\sum y^2-(\sum y)^2]}}$$

The correlation between sets of factors or variables measures the correlation between the factors.

Step 3: Data Mining Modelling

The study used the Waikato Environment for Knowledge Analysis (Weka) software for data mining. Weka is a collection of machine learning algorithms for data mining tasks (Frank *et al.,* 2016). In this study, logistic regression with a 10-fold cross-validation was performed using Weka. The study tested linear regression (Sohn *et al.,* 2007), neural networks (Sadatrasoul *et al.,* 2013) and SVMs (Sadatrasoul *et al.,* 2013). Weka is used to determine whether a data classification is true.

The process of data mining modelling used logistic regression, multilayer perceptron, SVM and receiver operating characteristics (ROC).

i. Logistic Regression

Logistic regression is a classification algorithm used to predict a binary outcome (1 / 0, Yes / No, True / False) of a set of independent variables (Agbemava *et al.,* 2016).

ii. Multilayer Perceptron

A multilayer perceptron is a feedforward ANN model that maps a set of input data onto a set of appropriate output. It uses a simple threshold or sigmoidal activation function to train algorithms for multi-layered networks. Each multivariate function can be represented in terms of a network diagram, such that there is a one-to-one correspondence between the model components and the elements of the diagram (Bishop, 1995). Units not treated as output units are called hidden units:

$$a_m = \sum_{i=1}^{q} E_{mh}^1 x_q + E_{m0}^1$$

where E_{mh}^1 is the weight of the first layer, going from input *h* to the hidden unit *m*, E_{m0}^1 denotes the bias of the hidden unit and *q* is the number of nodes. The output of the network is obtained by transforming the activation of the hidden units using a second layer of processing elements. A network is feed-forward if it is possible to attach successive numbers to the inputs to all the hidden and output units such that each unit only receive connections from inputs or units with a smaller number (Bishop, 1995).

iii. *Support-vector Machine*

The Sequential Minimal Optimisation algorithm is derived from the extremes of the decomposition method and optimising a minimal subset of just two points at each iteration (Platt, 1998).

iv. *Receiver Operating Characteristics*

A ROC area is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters (Fawcett, 2005).

The utilised confusion matrix consists of TP (True Positive), TN (True Negative), FP (False Positive) and FN (False Negative) values as shown in Table 1. Sensitivity is the calculation of the actual correctly identified instances, while Specificity is the calculation of correctly identified negatives (Vapnik, 2000).

Table 1: Table for the confusion matrix

	Prediction	
Actual	True Positive (TP)	True Negative (TN)
	False Positive (FP)	False Negative (FN)

$$Sensitivity = \frac{\text{Number of true positive}}{\text{Number of true positive} + \text{Number of false negative}}$$

$$Specificity = \frac{\text{Number of true negative}}{\text{Number of true negative} + \text{Number of false positive}}$$

The ROC graphs are two-dimensional graphs in which the t_p rate is plotted on the y-axis and f_p rate is plotted on the x-axis. A ROC graph depicts relative trade-offs between benefits (true positives) and costs (false positives).

Step 4: Credit Scoring

The weight of the independent variables obtained from the Weka logistic regression is used to construct the credit-scoring table.

Step 5: Post-modelling

Post-modelling is the monitoring and maintenance of the credit scoring model to ensure that the existing model remains reliable in the current business environment. Examples of post-modelling includes identifying changes in the characteristics of new loan applicants, changing the cut-off values for loan rejection or acceptance, determining whether the model produces the same credit scoring for loan applicants with the same risk factors.

Results

The results of this study are the step-by-step process of developing a credit scoring model for personal loans. This section demonstrates the outcome of the steps as outlined in the Research Methodology section.

Step 1: Data Massaging

The study used real personal loan customers' data supplied by one of the banks in Malaysia. The data supplied by the bank were from 2007 to 2018. However, only data between 2015 and 2018 were used to model the credit scoring. Data before 2005 were omitted as before that year, the bank was only a cooperative bank. Although an abundance of data was supplied, data before 2015 was incomplete and contained a significant number of errors. This was due to poor data management by the bank. The bank was initially a credit cooperative and was given the permission by Bank Negara to accept deposits in 2014. Credit cooperative is a cooperative licensed by Cooperative Commission Malaysia (Suruhajaya Koperasi Malaysia) to offer financing products to its member. Table 2 presents the data prior to the data massaging process.

Although the data supplied was from 2007, the data used for credit scoring modelling were from 2015 to 2018. The data in this period were deemed appropriate for the development of a credit scoring model as it aligned with the policy direction and business environment of the bank. In addition, based on the financial industry's best practice, credit scorecards were designed to predict the likelihood of future events over the next 12 to 24 months. The three steps in the data massaging process were executed on the data from the years 2015 to 2018.

Vintage Analysis

Vintage analysis is used to assess the creditworthiness of a loan portfolio by examining net charge-offs during the same origination period. The performance window is also determined using this analysis. The data was compiled in a vintage chart that listed each monthly vintage in rows, with the oldest at the top. Then, the chart was visualised as shown in Figure 2 to find the period (months of books), where the trend started to stabilise. In this case, the trend was assumed to stabilise starting from MOB18, which was equivalent to 18 months (the performance window).

Table 2: The shape of the data prior to the data massaging process

Details	Description
Period of data	2007 - 2018 (10 years)
Number of rows	107,226
Number of columns	206

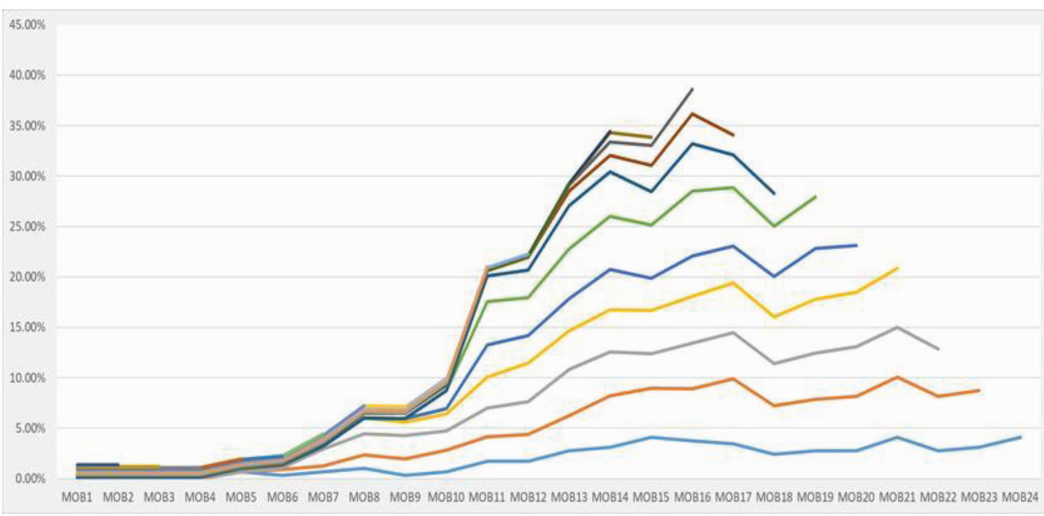


Figure 2: 90 days past due count by months on books

Roll Rate Analysis

Roll rates were the percentage of loans moving from one state of delinquency to another, for instance, from 30 days past due (DPD) to 60 DPD and to 90 DPD. The delinquency chart was constructed for each month and the percentage of backward roll rate was observed and presented in Table 3. The DPD with the lowest percentage of backward roll rate was taken as the definition of a bad account. In this case, 90+ DPD demonstrated the lowest backward roll rate (0%), indicating that there was a zero percentage that a default account will return (or rollback) to a good account.

Data Pruning

Data pruning was performed to address and resolve any anomaly in the data. The anomalies included missing data, outliers, illogical values and decoding the encoded values. After the pruning process, only 16 attributes (column) were chosen based on their suitability and relevance for the development of the credit scoring model, as shown in Table 4. The chosen attributes were considered as the key variables affecting the creditworthiness of a borrower.

The percentage of removed data are illustrated in Figure 3. The percentage of removed data (from pruning) were relatively low and imposed an insignificant effect on the overall modeling process.

The shape of the data after the data massaging process is summarised in Table 5.

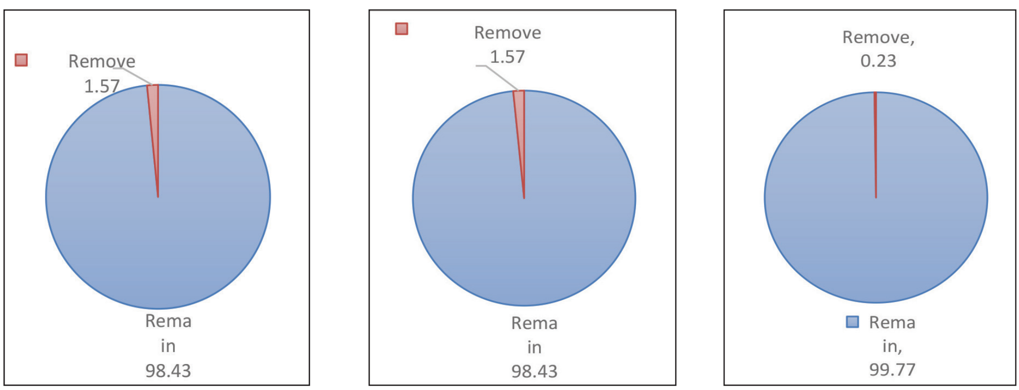
Finally, using the massaged data, the status of each account (bad/good) was determined based on the roll rate analysis. To recap, the roll rate analysis indicates that the appropriate period to define a bad account is 90 DPD or 3 months. Furthermore, an account that has been restructured is also considered a bad account. The proportion of good and bad accounts is illustrated in Figure 4.

Table 3: Delinquencies from April 2015 to May 2015

Delinquency (April 15 - May 15)		**Current Month**					
		Current (%)	**0 - 29 DPD (%)**	**30 - 59 DPD (%)**	**60 - 89 DPD (%)**	**90+ DPD (%)**	**Backward Roll Rate (%)**
Previous month	Current	47	53				
	0 - 29 DPD	3	54	43			3
	30 - 59 DPD	1	13	66	20		14
	60- 89 DPD	0	0	48	37	15	48
	90+ DPD	0	0	0	0	100	0

Table 4: List and description of attributes for the credit scoring model

Attributes	Description
Account number	The account number of the borrower
Age	Age of the borrower
Amount of finance	Amount of finance
Dependent	The number of borrower's dependent
Education	Borrower's education level
Gender	Sex of borrowers
Income	The borrower's monthly salary or income
Installment	Monthly payment
Monthly expenses	The borrower's monthly expenses
Net-worth	The amount of asset exceeding the liabilities of the borrower
Other commitment	Other loan commitments such as an education loan
Payment method	Type of payment, such as standing instruction, walk-ins
Payment schedule	Schedule for the monthly payment
Product code	Description not given by the bank
Period	Month of loan
Sector	Government-linked company (GLC), private or public employer of the borrower



Product Code - Loan other than personal financing

Product Code - Loan other than personal financing

Education Level - Blank

Figure 3: Percentage of removed data

Table 5: The shape of data after the data massaging process

Details	Description
Period of data	2015–2018 (3 years)
Number of rows	31,532
Number of columns	15+

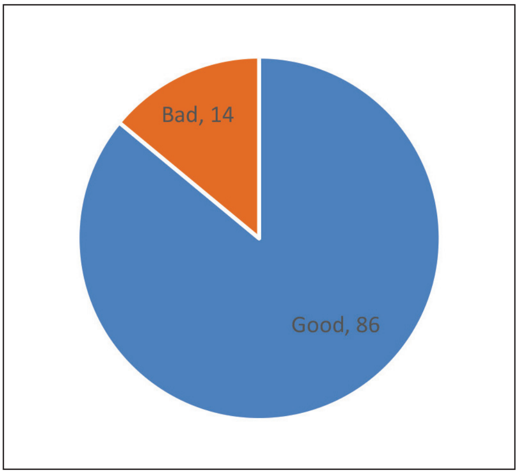


Figure 4: The proportion of good and bad accounts

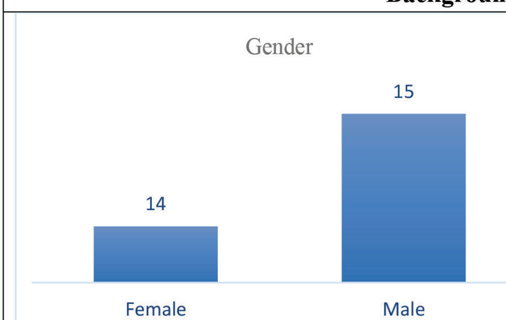
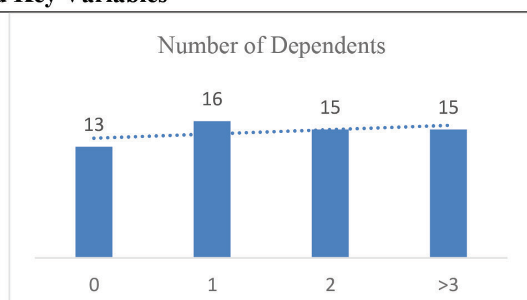
Step 2: Factor Analysis

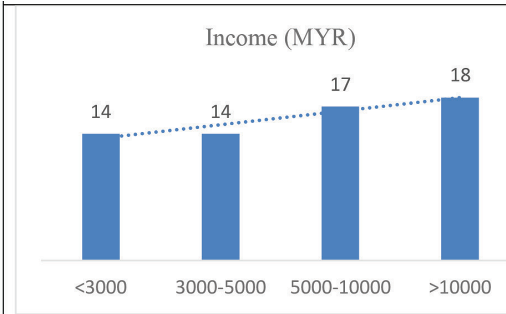
Factor analysis required the WoE, IV and correlation analysis between the variables.

Weight of Evidence

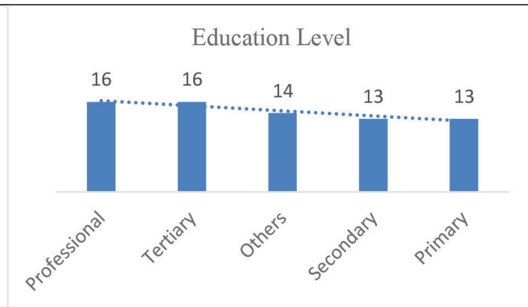
The WoE was computed to determine the strength of the relationship of each key variable (grouped in several bins) in influencing the account status. The value of the WoE for each bin was used to compute the overall IV for the particular variable. The rule of thumb was to find a decreasing or increasing trend of bad rate against each bin. The trend demonstrated significant influence on account status. Apart from that, the visualised percentage of bad rate retrieved from the computation of WoE for each key variable is summarised in Table 6. In Table 6, the key variables were separated into background and financial key variables. The visualisation of the percentage of bad rates helped in finding the logical and illogical trend of a key variable. In addition, the findings and conclusions on the respective key variable were also included.

Table 6: The visualisation of bad rates for each key variable

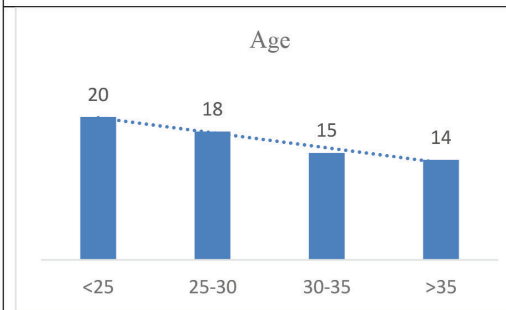
Background Key Variables	
 No significant difference in bad rate percentage across genders.	 No clear trend observed in bad rate percentage in terms of number of dependents.



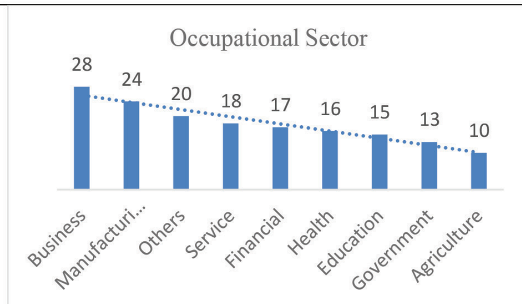
Illogical trend. As monthly income increased, the percentage of bad rate increased.



Illogical trend. As education level decreased, the percentage of bad rate decreased.

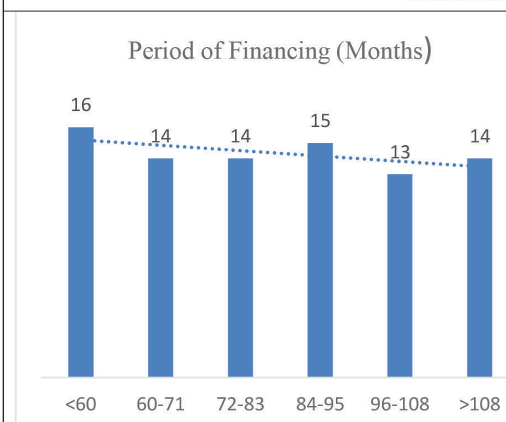


Logical trend. As age increased, the percentage of bad rate decreased.

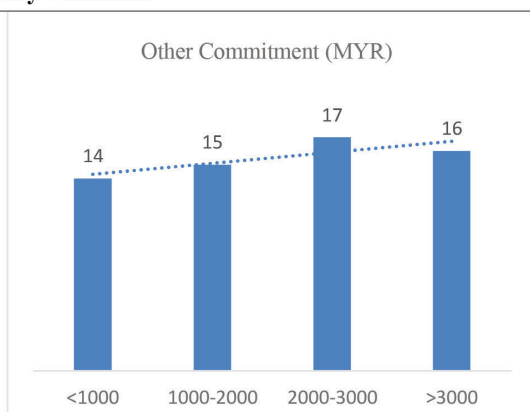


Clients from the business sector have the highest percentage of bad rate. The agricultural sector showed the least percentage of bad rates.

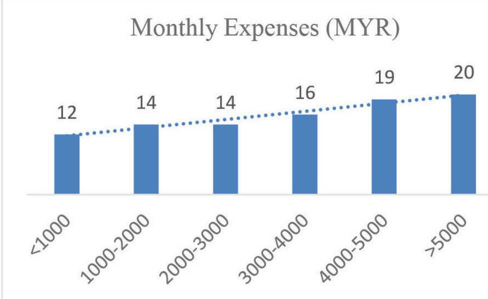
Financial Key Variables



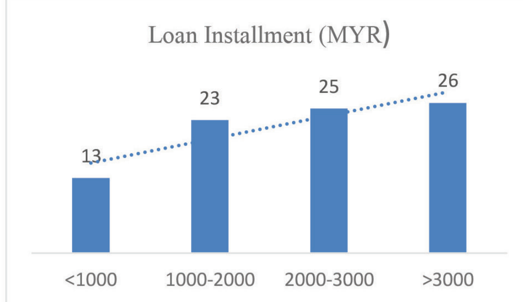
No observable trend.



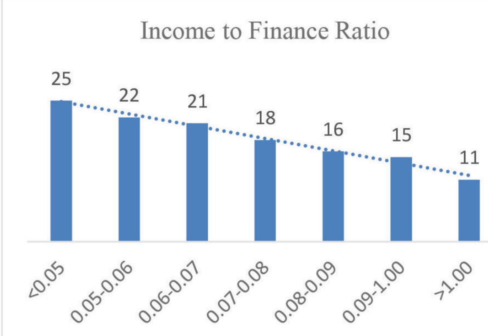
No observable trend.



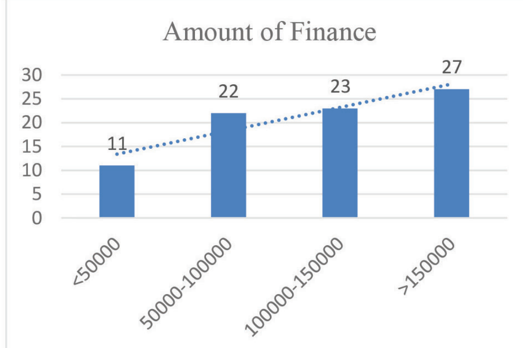
Logical trend. As monthly expenses increase, the percentage of bad rates increases.



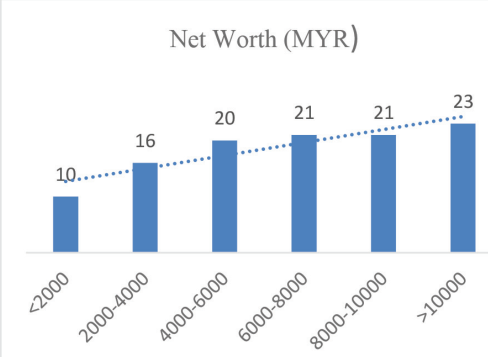
Logical trend. As installment increases, the percentage of bad rates increases.



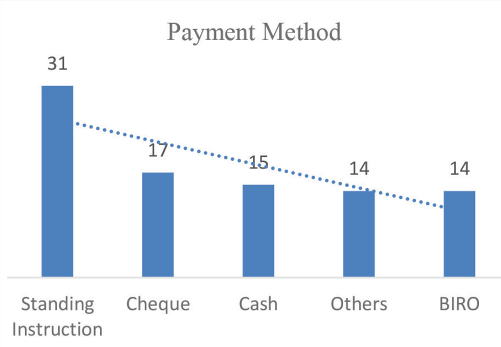
Logical trend. As the ratio increased, the percentage of bad rates decreased.



Logical trend. As the amount of finance increases, the percentage of bad rates increases.



Illogical trend. As the net worth of the client increases, the percentage of bad rate increases.



The account tends to be bad if the client opts to pay through standing instruction. The best payment method with the lowest probability of bad rate is BIRO (payment by salary deduction through the National Cooperative Union).

The study considered the income-to-finance ratio as one of the key variables as income demonstrated an illogical trend despite its importance in influencing the status of accounts. Hence, to ensure that income can be incorporated in the modelling process, the income-to-finance ratio was created and then observed to produce a logical trend.

Information Value

The value of WoE from each bin was used to calculate the IV. The IV determines the overall worthiness of a variable to be included as key variables influencing the account status. The IV for each key variable is summarised in Table 7. To define the relationship strength, the rule of thumb indicates that IV < 0.02 as unpredictive, $0.02 \leq IV < 0.1$ as weak, $0.1 \leq IV < 0.3$ as medium and $IV \geq 0.3$ as strong.

Even though the IVs provided numerical evidence of the worthiness of a variable to be accepted in the model, the final decisions were based on the industry's best practice. Although payment, income and age produced low IVs, the variables were accepted because according to the bank supplying the data, the variables were important in influencing the account status. To conclude, six out of 12 variables were chosen as the final key variables to be used in the development of the model.

Step 3: Data Mining Modelling

The formula for credit scoring is as follows:

$$Score = -(WOE \times \beta + \alpha \div 1) \times factor + (\frac{offset}{1})$$

$$\beta = Linear\ Regression\ Bad\ data\ from\ Weka\ show\ in\ Table\ 2$$

$$\alpha = Linear\ Regression\ Bad\ data\ from\ Weka\ show\ in\ Table\ 2\ from\ Intercept$$

The logistic regression produced the coefficients for each variable as shown in Table 8.

$$factor = PDO/\ln 2$$

$$Offset = Point - [factor \times \ln(odds)]$$

*Assume PDO = 20; Odds = 50 and Point = 600

Table 7: Summary of the Information Value for key variables

Attributes	IV	Strength	Verdict
Income-finance ratio	0.106202219	Medium	Accepted
Net worth	0.102945235	Medium	Rejected - Illogical trend
Installment	0.058149089	Weak	Accepted
Sector	0.025483978	Weak	Accepted
Monthly expenses	0.024760641	Weak	Accepted
Payment	0.016927813	Unpredictive	Accepted
Education level	0.008109625	Unpredictive	Rejected - Illogical trend
Dependent	0.006488738	Unpredictive	Rejected - Illogical trend
Age	0.006341824	Unpredictive	Accepted
Other commitment	0.00409305	Unpredictive	Rejected - too low
Period of finance	0.001276868	Unpredictive	Rejected - too low
Gender	0.000372095	Unpredictive	Rejected - too low

Table 8: Coefficients from logistic regression modelling

Variable	Definition	Class (Bad)	Odds Ratio
INST_LE	Payment every month	-0.1857	1.2041
AGE_LE	Age of the borrower	-0.0489	0.9522
EXP_LE	Monthly expenses	-0.1833	0.8325
SECTOR_LE	Government-linked company (GLC), private or public company of the borrower's employee	-0.0559	0.9456
PAY_LE	Type of payment such as standing instruction	-0.1497	0.861
INCFINR_LE	Income-to-financial ratio of borrowers	-0.2045	0.815
Intercept		2.2359	

Comparison for linear regression, neural networks and SVM is presented in Table 9.

Table 9: Comparison between linear regression, neural network and support-vector machine

Method	Correctly Classified	Incorrectly Classified	ROC Area
Linear regression	61.5%	38.5%	0.64
Multilayer perceptron	61.7%	38.3%	0.63
Support-vector machine	61.0%	39.0%	0.58

Although multilayer perceptron had only 0.2% correctly classified linear regression, the linear regression was accepted as the ROC area was higher than the multilayer perceptron. In general, an acceptable model is produced with a 61.5% accuracy and an ROC area of 0.64.

Step 4: Credit Scoring

The coefficients obtained from logistic regression were used to construct a scoring table. The scoring table facilitated the business decision-making process, especially the evaluation of the creditworthiness of new clients.

Cut-off Analysis

The cut-off analysis provides leverage to a bank to set an acceptance tolerance that is in line with the current policy's direction and business environment. The interactive cut-off analysis is shown in Figure 5. A higher setting of cut-off point reflects that the bank is tolerable in accepting loan applications from new clients. Nevertheless, high cut-off points come with a high risk of possible loan defaulters in the future. The trade-off between profit and risk reflected via cut-off analysis is subject to the discretion of the bank.

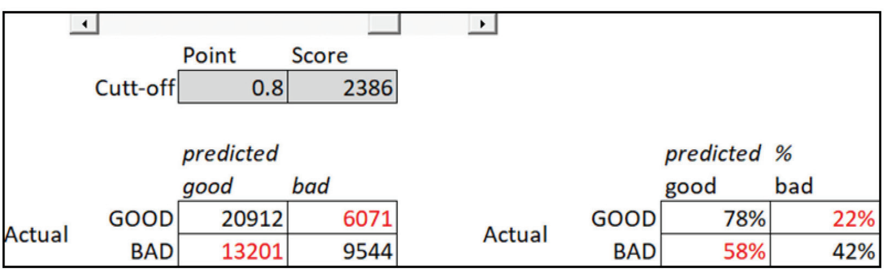


Figure 5: Interactive cut-off analysis

Step 5: Post-modelling

Post-modelling maintenance and monitoring are as important as developing the model to ensure that the existing model remains reliable in the current business environment. The rule of thumb for this is to re-valuate and revise the model every six months. If the bank finds major changes in the model performance, the study proposes a rebuild of the credit scoring model. On the other hand, minor changes in the model performance may be ignored or put under observation.

Discussion

A five-step credit scoring model using the seven-step credit scoring model (Siddiqi, 2012) as reference using actual data supplied by a bank was developed in this study. The model was tested using dataset of real personal loan customers. The five-step model was sufficient for determining the credit scoring worthiness of loan applicants. Banks may have an abundance of data. However, not all of the data were useful. During the model development phase, there was a data quality issue. However, the issue was resolved using data massaging. Although automated tools such as a credit-scoring model are useful in handling large amounts of data and to performcomplex calculations rapidly, humans still undertake the decision-making process (Sum, R. M., 2018). During the development process, the variables payment, income and age produced low IVs. However, the bank requested that the variables were accepted because the variables were important in influencing the account status.

Conclusion and Future Direction

A personal loan credit scoring model using data mining techniques was developed in this study, and the steps in developing the particular model were discussed in detail. The model enabled loan officers to automate the decision-making process and produce an accurate prediction of the creditworthiness of applicants. The study acknowledged the omission of several variables in the development of the credit scoring model. The variables were net worth, education level, number of dependents, other commitments, period of financing and gender. Future incorporation of these variables may have the potential to improve the accuracy and predictive power of the model. Furthermore, the accuracy of the model can be increased by using it on datasets from other banks. Finally, the importance of model maintenance is stressed to ensure the model's continuous reliability in the current business environment.

Acknowledgements

The authors would like to acknowledge the supporters of this project, which is in part by the Fundamental Research Grant Scheme FRGS, Ministry of Higher Education, with Application ID 225085-266898 and reference code - FRGS/ 1/2018/ICT02/USIM/02/1, Universiti Sains Islam Malaysia (USIM), Faculty of Science and Technology (FST), Project code (USIM/FRGS/ FST/055002/51918). Thank you to Mr. Zabid, Mr. Khalili and Mr. Zulhakim for contribution to data and analysis.

References

Abdou, H. A., & Pointon, J. (2011). Credit scoring, statistical techniques and evaluation criteria: A review of the literature. *Intelligent Systems in Accounting, Finance and Management*, *18*(2-3), 59-88.

Agbemava, E., Nyarko, I. K., Adade, T. C., & Bediako, A. K. (2016). Logistic Regression Analysis of predictors of loan defaults by customers of non-traditional banks in Ghana. *European Scientific Journal*, *12*(1), 175-189.

Ala'raj, M., Abbod, M., & Radi, M. (2018). The applicability of credit scoring models in emerging economies: An evidence from Jordan. *International Journal of Islamic and Middle Eastern Finance and Management*, *11*(4), 608-630.

Arya, S., Eckel, C., & Wichman, C. (2013). Anatomy of the credit score. *Journal of Economic Behavior & Organization*, *95*, 175-185.

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*, *54*, 627-635.

Bequé, A., & Lessmann, S. (2017). Extreme learning machines for credit scoring: An empirical evaluation. *Expert Systems with Applications*, *86*, 42-53.

Bishop, C. M. (1995). *Neural network for pattern recognition*. Oxford: Clarendon Press.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321-357.

Chopra, A., & Bhilare, P. (2018). Application of ensemble models in credit scoring models. *Business Perspectives and Research*, *6*(2), 129-141.

Crook, J. N., Edelman, D. B., & Thomas, L. C. (2007). Recent developments in consumer credit risk assessment. *European Journal of Operational Research*, *183*, 1447-1465.

Dong, G., Lai, K. K., & Yen, J. (2010). Credit scorecard based on logistic regression with random coefficients. *Procedia Computer Science*, *1*(1), 2463-2468.

Edelman, D. B., & Crook, J. N. (2002). *Credit scoring and its applications.* Philadelphia: Society for Industrial Mathematics (SIAM).

Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(2006), 861-874.

Frank, E., Hall, M. A., & Witten, I. H. (2016). *The WEKA Workbench. Online Appendix for Data Mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.

Ghosh, S. (2019). Loan delinquency in banking systems: How effective are credit reporting systems? *Research in International Business and Finance*, *47*, 220-236.

Hand, D. J., & Henley, W. E. (1997). Statistical classification models in consumer credit scoring: A review. *Journal of the Royal Statistical Society: Series A (General)*, *160*, 523-541.

Hsieh, N. C. (2004a). An integrated data mining and behavioral scoring model for analyzing bank customers. *Expert Systems with Applications*, *27*(4), 623-633.

Hsieh, N. C. (2004b). Hybrid mining approach in the design of credit scoring models. *Expert Systems with Applications*, *28*(4), 655–665.

Hui, L., Li, S., & Zongfang, Z. (2013). The model and empirical research of application scoring based on data mining methods. *Procedia Computer Science*, *17*, 911-918.

Kamleitner, B., & Kirchler, W. (2007). Consumer credit use: A process model and literature review. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *57*(4), 267-283.

Kim, Y. S., & Sohn, S. Y. (2007). Technology scoring model considering rejected applicants and effect of reject inference. *Journal of the Operational Research Society*, *58*(10), 1341-1347.

Kruppa, J., Schwarz, A., Arminger, G., & Ziegler, A. (2013). Consumer credit risk: Individual probability estimates using machine learning. *Expert Systems with Applications*, *40*(13), 5125-5131.

Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247*(1), 124-136. Doi: 10.1016/j.ejor.2015.05030.

Miller, S. (2015). Information and default in consumer credit markets: Evidence from a natural experiment. *Journal of Financial Intermediation*, *24*(1), 45-70.

Pentzlin, D. (2011). *Why to integrate sustainability criteria in financial regulation. Friends of the Earth*. URL: www. foeeurope. org.

Platt, J. (1998) “Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines” Microsoft.com MSR-TR-98-14 | April 1998. https://www.microsoft.com/enus/research/publication/sequential-minimal-optimization-a-fast-algorithm-for-training support-vector-machines/

Salem, A. S., & Mahamad, S. (2020). *PITEH: Providing Financial Identities to Those without Credit Score.* 2020 International Conference onComputational Intelligence (ICCI). Doi:10.1109/icci51257. 2020.9247 779

Sadatrasoul, S. M., Gholamian, M. R., Siami, M., & Hajimohammadi, Z. (2013). Credit scoring in banks and financial institutions via data mining techniques: A literature review. *Journal of AI and Data Mining*, *1*(2), 119 -129.

Siddiqi, N. (2012). *Credit risk scorecards: Developing and implementing intelligent credit scoring* (Vol. 3). Hoboken, New Jersey: John Wiley & Sons.

Sohn, S. Y., & Kim, H. S. (2007). Random effects logistic regression model for default prediction of technology credit guarantee fund. *European Journal of Operational Research*, *183*(1), 427-478.

Sohn, S. Y., Kim, H. S., & Moon, T. H. (2007). Predicting the financial performance index of technology fund for SME using structural equation model. *Expert Systems with Applications*, *32*(3), 890-898.

Sohn, S. Y., & Kim, Y. S. (2012). Behavioral credit scoring model for technology-based firms that considers uncertain financial ratios obtained from relationship banking. *Small Business Economics*, *41*(4), 931-943.

Sum, R. M. (2015). *Risk Prioritisation (RP): A decision making tool for risk management*. (PhD thesis) Macquarie University, Sydney.

Sum, R. M. (2018). Using Mathematics to quantify subjective decisions: Application of analytic hierarchy process to risk assessment. *Journal of Advanced Research Design*, *44*(1), 7-19.

Tan, P. N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. Wesley, Boston: Pearson Addison.

Vapnik, V. N. (2000). *The nature of statistical learning theory*. Springer Verlag.

Ziemba, P., Zalas, A. R., & Becker, J. (2020). Client evaluation decision model in the credit scoring tasks. *Procedia Computer Science*, *176*, 3301-3309.