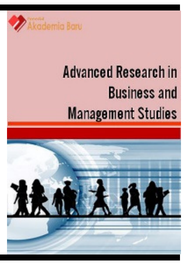
Penerbit Akademia Baru

Journal of Advanced Research in Business and Management Studies

Journal homepage: www.akademiabaru.com/arbms.html

ISSN: 2462-1935



Credit scoring models: Techniques and issues

Open Access

Yosi Lizar Eddy [1], Engku Muhammad Nazri Engku Abu Bakar [*,2]

[1] Risk Quantification Section, Risk Management Department, Bank Muamalat Malaysia Berhad, 21 Jalan Melaka, 50100 Kuala Lumpur, Malaysia
[2] Decision Science Department, School of Quantitative Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia

ARTICLE INFO

Article history:
Received 30 January 2017
Received in revised form 22 May 2017
Accepted 4 June 2017
Available online 6 June 2017

Keywords:
Credit scoring model, Logistic regression, DEMATEL, Analytic hierarchy process

ABSTRACT

This paper presents a brief review on the current available techniques for credit scoring model, namely the statistical-based models and the artificial intelligence/machine learning- based models. It is then followed by the suggestions on how to revise the credit scoring model that is currently being adopted by any credit risk management, if revision is needed. The revision of the model involves the selection of criteria to be included as well as the weights to be given for the criteria. Some potential techniques in selecting the criteria and determining the weights for the selected criteria are also discussed.

1. Introduction

Giving out loan is an insecure business but at the same time, it is one of the major sources of income to most banks. Banks would prefer not to allow credit to those customers lacking capacity to pay back the credit given. Be that as it may, after some time, a certain percentage of the credits will eventually transform into bad loan regardless of the possibility that the banks tighten its credit policy [1]. Analyzing the non-performing loans (NPLs) data will effectively measure the quality of credit endorsement process. The loan granting process must be observed vigilantly and banks should formulate an effective credit risk management.

In most banking institutions, loan application processes are done on a centralized approach by a department which assesses and segregates financing applications between risky and non-risky customers based on the banks' credit scoring model. This credit scoring process will determine who should get credit and how much credit should be granted with the intention to minimize the risk of loan losses and delinquency rate due to the costly misclassifying error. Thus, the bigger the misclassifying error created, the greater the risk produced [1-3]. Despite the thorough process and regardless of the possibility that the banks tighten its credit policy, after some time, a certain

[*] *Corresponding author.*
E-mail address: enazri@uum.edu.my (Engku Muhammad Nazri Engku Abu Bakar)

percentage of the credits will eventually transform into bad loans. This raises a question of whether the credit scoring model was constructed efficiently, particularly with respect to the selection of relevant factors/variables for the model and the weights assigned to those variables in the model. Therefore, there is a need for the bank management to review its current loan application screening approach, i.e. the existing credit scoring model. Should the current model be maintained? Should the criteria used in the model be revised? Is the weight attached to each criterion in the model appropriate? Are there any other better and simpler approaches that can be utilized?

The objectives of this paper are therefore:
i. To review the available credit scoring model by identifying the techniques, the pros and the cons.
ii. To list the suitable criteria to be included in a credit scoring model.
iii. To put forward some recommendation on the selection of the appropriate model to be adopted.

2. Credit Scoring Techniques

Credit scoring techniques can be divided into two different methods namely statistical-based method and artificial intelligence/machine learning-based method [4-5].

2.1 Statistical-based Credit Scoring Models

There are various statistical-based credit scoring model that have been introduced such as linear discriminant analysis [6], decision trees [11-12], Markov chain analysis [13-14], probit analysis [10] and logistic regression [15-16].

The technique of linear discriminant analysis assumes that there are two populations of individuals, denoted as `1' for defaulters and `0' for non-defaulters; each is characterized by a multivariate distribution of a set of attributes, X , including factors such as age, income, family size, credit history, occupation, and so on [52]. The formula for the linear discriminant analysis is as follows [8]:

$$Z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_n X_n \quad (1)$$

where Z symbolizes the discriminant Z-score, α is the intercept term, $\beta_1 \ldots \beta_n$ stand for coefficients in the linear combination of the explanatory variables for Xi for i = 1 … n.

Linear discriminant analysis is one of the earliest common traditional statistical techniques used for constructing credit scoring models. However, this technique requires rather restrictive statistical assumptions that are seldom satisfied in real life.

Classification or decision trees on the other hand, can be used to classify a potential loan applicant to a predefined set of rating groups. It is a flowchart-like structure in which each internal node represents an analysis on an attribute, each branch represents the outcome of the test and each leaf node represents a class label (decision taken after computing all attributes). The applicant is firstly classified according to the most relevant variable, then into subgroups according to other variable, and so on [53]. In this situation, the paths from root to leaf represent classification rules. The system considers all possible splits to find the best one, and the winning sub-tree is selected based on its overall error rate or lowest cost of misclassification. In other words, the decision tree credit scoring model will find a classifier that separates the good credit sample from the bad credit sample. The algorithm begins with a root node containing a sample of good and bad credit applicants. Then, the algorithm loops over all possible binary splits in order to find the attribute x and corresponding cut-off value which gives the best separation into one side having mostly good credits and the other mostly bad credits [11].

Meanwhile, Markov Chain, also known as transition matrix, is a mathematical model which defines the probability of a borrower moving from one state to other states. Depending on the data available the matrix is developed by having each entry in the matrix to represent the probability that the borrower will move to state j given that it starts from state i (usually defined as the time to transit from state which in this case can be one hour, one day, one month, or even longer) as shown by the matrix below [54].

States	**A(1)**	**A(2)**	**. . . . .**	**A(N-1)**	**A(N)**
A(1)	P(1,1)	P(1,2)	. . . . .	P(1,N-1)	P(1,N)
A(2)	P(2,1)	P(2,2)		P(2,N-1)	P(2,N)
.	.	.	. . . . .	.	.
.	.	.	. . . . .	.	.
.	.	.	. . . . .	.	.
.	.	.	. . . . .	.	.
.	.	.	. . . . .	.	.
A(N-1)	P(N-1,1)	P(N-1,2)	. . . . .	P(N-1,N-1)	P(N-1,N)
A(N)	P(N,1)	P(N,2)	. . . . .	P(N,N-1)	P(N,N)

Next, with some simple mathematical formulation, the transition matrix will be converted into steady-state, or long-run probability for a borrower to be a potential good client or a potential bad client.

Finally, logit and probit regression analysis are the multivariate techniques which allow for estimating the probability that an event occurs or not, by predicting a binary dependent outcome from a set of independent variables [55]. The response, y_i, is equal to 0 if default occurs (with probability Pi) and to 1 if default does not occur (with probability $1 – P_i$). In the regression models, we wish to model the probability P_i that the default will occur by specifying the model

$$P_i = f(\alpha + \beta' x_i) \tag{2}$$

where xi are particular financial indicators and α, β are estimated parameters. In case of probit model we use the cumulative distribution function of normal distribution:

$$P_i = \int_{-\infty}^{\alpha+\beta_i} \frac{1}{\sqrt{2\Pi}} \exp\left(-\frac{1}{2}t^2\right) dt \tag{3}$$

Due to nonlinear features of these models it is necessary to use maximum likelihood method to estimate the parameters. Given Pi and assuming that defaults are independent, we can form the logarithm of likelihood function as follows:

$$\ln L = \sum_{i=1}^{n} y_i \ln P_i + \sum_{i=1}^{n} (1 - y_i) \ln(1 - P_i) \tag{4}$$

In non-payment predictions, Neophytou, Charitou *et al.* [17] suggested that the logistic regression method is better in comparison to other methods. This statement is further supported by Bolton [15] who suggested that credit scoring model using logistic regression is easy to elaborate and executed and this technique has been widely being acknowledged as method of choice by the banking industry. In addition, Gardner and Mills [18] suggested that logistic regression can be used to calculate the probability of default to identify the severity of the delinquency problem and consecutively construct a suitable action to the problem so that delinquent accounts do not necessarily end up as a defaulted loan accounts.

Logistic regression, also called logit regression or logit model is being used for forecasting the result of a categorical dependent variable (DV) in view of one or more independent variables (IVs). It is applied as a part of assessing the parameters of a qualitative model of which the DV is either binary for two categories of DV or multinomial if there are more than two categories of DV [16]. Logit regression estimates the connection between the categorical DV with IVs with probability scores as the predictive value of the DV. Its function is similar to probit regression except that probit regression acts as standard distribution function instead of logistic function.

The formula for logistic regression is as follows:

$$\text{Logit}(p) = \log\,[p/\,(1 - p)] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots.. + \beta_n X_n \quad (5)$$

where,

p represents the outcome probability,
α represents the intercept term,
$\beta_1 \ldots \beta_n$ correspond with the coefficients in the linear grouping of descriptive variables,
X_i, for $i = 1 \ldots n$.

2.2 Artificial Intelligence-based/Machine Learning Methods

Some of the AI-based methods being suggested and explored by researchers are artificial neural networks [19-22], genetic algorithms [23-25], and artificial immune system [26-27]. Upon identifying the set of factors to be used, the weight for each factor will be set randomly. Then the AI-approach is used to generate the best combination of weights that can produce the highest percentage of correct prediction.

Artificial Neural Networks are computational techniques that present a mathematical model based upon the neural structure of intelligent organisms and who acquire knowledge through experience. Introduced in the eighties, an artificial neural network processes certain characteristics and produces replies like those of the human brain. It is developed using mathematical models in which the following suppositions are made [59]:

i. Process of information takes place within the so-called neurons;
ii. Stimuli are transmitted by the neurons through connections;
iii. Each connection is associated to a weight which, in a standard neural network, multiplies itself upon receiving a stimulus;
iv. Each neuron contributes for the activation function (in general not linear) to determine the output stimulus (response of the network).

Genetic algorithm (GA), resembling the evolution of the species proposed by Darwin [60] comprises of a population of chromosomes representing the various possible solutions for the proposed problem. Solutions that are selected to shape new solutions via a cross-over process are determined according to the fitness of the parent chromosomes. Thus, the fitter the chromosome is, the higher the possibility of reproducing itself. This process is repeated until the rule of halt is satisfied. Every genetic algorithm goes through the following stages:

Step 1: Initially a population is generated, formed by a random set of individuals (chromosomes) that may be viewed as possible solutions for the problem.
Step 2: A function of fitness is defined to evaluate the "quality" of each one of the chromosomes. Based on the results of the fitness function, a percentage of the best fit is maintained while the others are rejected (Darwinism).
Step 3: Two best-fit parents are chosen for cross-over process to produce a pair of offspring based on a specific cross-over criterion. The same criterion is used with another chromosome and the material of both chromosomes is exchanged. If there is no cross-over, the offspring is an exact copy of the parents.
Step 4: Mutation, which is an alteration in one of the genes of the chromosome, may be performed next. The purpose of mutation is to avoid that the population converges to a local maximum. Thus, should this convergence take place, mutation ensures that the population will jump over the minimum local point, endeavoring to reach other maximum points.
Step 5: Verification of the halt criterion: once a new generation is created, the criterion of halt is verified and should this criterion not have been met, one returns to the stage of the fitness function (Step 2).

In a conventional artificial immune system (AIS) algorithm, a classifier system is constructed as a set of exemplars that can be used to classify a wide range of data and in the context of immunology, the exemplars are known as B-cells and the data to be classified are known as antigens. A typical AIS algorithm operates as follows [61]:

i. First, a set of training data (antigens) is loaded and an initial classifier system is created as a pool of B-cells with attributes either initialised from random values or values taken from random samples of antigens.
ii. Next, for each antigen in the training set, the B-cells in the cell pool are stimulated. The most highly stimulated B-cell is cloned and mutated, and the best mutant is inserted in the cell pool. To prevent the cell pool from growing to huge proportions, B-cells that are similar to each other and those with the least stimulation levels are removed from the cell pool.
iii. The final B-cell pool represents the classifier.

Machine learning on the other hand, refers to a set of algorithms specifically designed to tackle computationally intensive pattern-recognition problems in extremely large datasets. These techniques include radial basis functions, tree-based classifiers, and support-vector machines, and are ideally suited for consumer credit-risk analytics because of the large sample sizes and the complexity of the possible relationships among consumer transactions and characteristics. Having identified the pattern, the forecasting model will be developed using any suitable forecasting techniques. Li *et al*. [28] and Bellotti and Crook [29] discussed in detail some of the applications of machine learning-based model to consumer credit.

3. Method of Choice

Logistic regression is superior to other methods in predicting defaults ([4], [15], [17]). It is easy to explain, very tractable, convenient, most practical and favorable technique in practice. However, Neophytou *et al*. [17] suggested that neural network is equally superior as its overall predictive ability is comparatively high. Nevertheless, logit model produces slightly lower type I error rates i.e. error in classifying bad loan as a good loan with an average of score of 16% compared to 17% for the neural networks. In addition, Koh *et al*. [8] noted that logistic regression is the most stable model despite of neural network having the best overall accuracy rate and Vojtek and Kočenda [4] suggested that even though the AI-based methods such as neural networks able to perform with missing values and multicollinearity issues, the processes are mathematically demanding and some techniques are complicated to elaborate. Finally, even though neural network can possibly create fantastic results, the failure to give details of the outcome show a real disadvantage to put these techniques into practice.

Regardless of which technique is being implemented currently, to improve the accuracy of the prediction, i.e. to revise the current credit scoring model used, we suggest the following:
For the case, whereby all the required historical data are available,

i. Reformulate the currently-used model using the historical data available; or
ii. Using the same historical data, calculate the probability of default for each type of clients.

Then, set the benchmark probability for loan acceptance/rejection. The benchmark probability can be set using either the expected profit margin, or expected loss. This way, the loan granting process can be done using a simple probability screening process without having to deal with a complicated scoring model.

On the other hand, in the case of the non-availability of historical data, particularly when we want to revise the model by including some new variables into the model, two possible techniques to be used to identify the suitable set of criteria will be the Decision-Making Trial and Evaluation Laboratory (DEMATEL) technique and the Analytic Hierarchy Process (AHP) technique. The way both techniques work is similar to the way the initial solution for the AI-techniques is produced. However, rather than allowing the AI-techniques to search for the best combination of weights, DEMATEL and AHP generate the weights through expert opinion and evaluation.

The next three sections will discuss on the potential factors that can be included in the credit scoring model, followed by the explanation on DEMATEL and AHP, respectively.

4. Factors to be considered in Credit Scoring

The objective of credit scoring is to estimate, by experimental way, the risk of giving out loan to specific criteria of borrower. This objective can be attained with a decision model that enables a bank to forecast future repayment trend of a candidate with identical attributes to historical data of past applicants. Variables containing the information of the characteristics of borrowers are applied into the credit scoring model. Credit scoring model will utilize this data to forecast the likelihood that loan will be promptly reimbursed ([1], [33]). It is therefore important to develop a scoring model which consists of the right variables. The criteria or variables or factors of borrowers with problem debts can be segregated based on demography and education; events in life; financial literacy and over indebtedness; credit lending policy and the role of government [34 - 37]. These factors can be further segregated into two common features [4-5]: reliability in estimating the probability of default and the ability to provide explanations when a method in credit-scoring is being applied to evaluate the loan submission. The factors can also be separated into four different indicators as the following [5]:

i. Financial Indicators - financial status or position of the loan applicant in repaying the loan.
ii. Demographic Indicators - quantifiable characteristics such as such as age, race, sex, and level of education, among others.
iii. Employment Indicators - employed, self-employed, or unemployed.
iv. Behavioral Indicators - history with financial institutions.

Besides all the factors mentioned above, some of these other factors can be considered as well. The factors are:

i. Life Events - problematic debt was more often due to a gradual accumulation of circumstances and range of reasons but the main ones were redundancy or job loss, followed by relationship breakdown, and sickness or disability, which usually result in late payments or defaults ([34], [39–40]).
ii. Credit Lending Practices - for some people credit had been too easy to obtain. Cross selling was often encouraged, by 'rewarding' customers' ability to repay by offering a further loan. Again, the 'manageability' of installment amounts, and the flexibility for borrowers to miss repayments from time to time was seen as more important than the interest being charged. Banks' demands for up-front lump-sums to start repayment plans reduced their aptitude to cope with their debts more effectively. In addition, excessive bank charges and penalties exacerbated debts, make it even more difficult for borrowers to meet their ends and service their existing debts [39].
iii. Government Policy - Ernst & Young, in the World Islamic Banking Competitiveness Report 2013–14 reported that structurally flawed economic policies that provided little or no incentives for long term savings may lead to maturity mismatch situation like what happened during 2007-08 financial crises whereby banks have significant long term asset but lack of short-term liability such as savings. For instance, government policy in some countries focused on rising home ownership through various programs such as providing down payment assistance to help low-income and minority families obtain mortgages which led to the root of crisis of large increase of sub-prime mortgage lending [41].

5. Decision-Making Trial and Evaluation

Decision-making trial and evaluation laboratory technique (DEMATEL) can be used to identify the causal-effect relations between the candidate criteria to be included in the credit scoring/risk model. The steps to apply DEMATEL can be simplified as follows [30-32]. Firstly, each decision maker is requested to specify the direct influence between any two criteria based on a scale consisting of 0,1,2,3, and 4 representing "no influence", "low influence", "medium influence", "high influence", and "very high influence", respectively. Then, execute steps [56]:

i. Calculate the direct-influenced matrix normalization.
ii. Produce the total-relation matrix.
iii. Draw the causal diagram.
iv. Obtain the inner dependence matrix and impact relationship map.
v. Obtain the inner dependence matrix. In this step, the sum of each column in total-relation n×n matrix is equal to 1 by the normalization method and then the inner dependence matrix can be acquired.

Step 1: Calculate the average matrix Z = [zij]. A group of m experts and n factors are used in this step. Each expert is asked to view the degree of direct influence between two factors using the scales given earlier based on pair-wise comparison. The degree to which the expert perceived factor i to affect factor j is denoted as xij. For each expert, an n x n non-negative matrix is constructed, where k is the number of the experts participating in the evaluation process with 1≤ k ≤ m. Thus, X1, X2, X3,.., Xm are the matrices from m experts. The average matrix Z = [zij] is obtained by calculating each entry as:

$$z_{ij} = \frac{1}{m}\sum_{i=1}^{m} x_{ij}^{k} \tag{6}$$

Step 2: Calculate the normalized initial direct-relation matrix D = [dij]. Each element in matrix D is ranged between [0,1]. It is calculated as:

$$D = \lambda Z \tag{7}$$

where

$$\lambda = \min\left[\frac{1}{\max(i)\sum_{j=1}^{n}|z_{ij}|}, \frac{1}{\max(j)\sum_{i=1}^{n}|z_{ij}|}\right] \tag{8}$$

Step 3: Derive the total relation matrix T. The full direct/indirect influence matrix T is derived using the following formula:

$$T = D(I - D)^{-1} \tag{9}$$

where *I* is an *n* x *n* identity matrix.

Step 4: Calculate the sum of rows and columns of matrix T. In this total-influence matrix T, the sum of rows and columns are represented by vector r and c, respectively.

$$r = [r_i]_{nx1} = (\sum_{j=1}^{n} t_{ij})_{nx1} \tag{10}$$

$$c = [c_j]_{1xn}^{T} = [\sum_{j=1}^{n} t_{ij}]_{1xn}^{T} \tag{11}$$

r_i indicates the direct and indirect effects that factor i has on the other factors. c_j on the other hand shows the sum of direct and indirect effects that factor j has received from the other factors. If $j = i$, the value of ($r_i + c_i$) represents an index of the strength of influences given and received or the degree of the central role that factor i plays in the problem. In contrast, the value of ($r_i - c_i$) shows the net contribution by factor i on the system. If ($r_i - c_i$) is positive, factor i is affecting other factors. When ($r_i - c_i$) is negative, factor *i* is being influenced by other factors [56 – 57].

Step 5: Set a threshold value α. α is computed by the average of the elements in matrix *T* to eliminate some minor effects in *T* [58].

$$\alpha = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}[t_{ij}]}{N} \tag{12}$$

where *N* is the total number of elements in matrix T.

Step 6: Build a cause and effect relationship diagram. The cause and effect diagram is constructed by mapping all coordinate sets of ($r_i + c_i$, $r_i - c_i$) to visualize the complex interrelationship and provide information to judge which are the most important factors and how influence affected factors [30]. The factors where tij is greater than α will be selected and shown in the cause and effect diagram [58], thus becoming the potential variables or factors to be included in the scoring model. The selected variables will then be given proper weights and the weights will be decided with the help from the experts.

6. Determining the Weight of Criteria using UHP

The primary objective in credit scoring is to develop an effective scoring model which contains only a small number of predictor variables. These scoring algorithms or "scorecards" are then used to evaluate all credit applicants in the future. This allows for consistency in credit evaluation and efficiency in processing. When implemented effectively, the scorecard should be able to rank order the entire population of applicants by risk [42]. Having selected the predictor variables, the proper weights to indicate the level of influence each variable has on the final score in the credit risk model must be determined. Among the techniques to determine weights are weight-of-evidence and information value technique [43], ELECTRE [44] and analytic hierarchy process (AHP) [45]. Among those techniques, AHP is the most widely used technique.

AHP is a technique that simplifies a complex problem by means of hierarchical analysis methodology, which enables subjective judgments among different criteria. AHP has indeed also been used to build a credit risk assessment model [46]. Serrano-cinca *et al*. [47] claimed that AHP provides a good solution to deal with the financial vs. social problem. The model assesses the credit history of the applicant (past), accounting information and intangible assets from the applicant itself (present), and the project to be financed, from the financial and from the social point of view (future). These criteria are reflected in different measurable indicators, which are evaluated by credit analysts. Beyond a score, the model allows identifying the strengths and weaknesses of the project to be financed.

AHP, as a tool to build expert systems, allows incorporating the knowledge of human specialists in a given subject into computer software that allowed the calculation of financial ratios or discounted cash-flows and the matrix calculus needed by AHP. The software has four main tabs, representing each of the AHP stages: (1) modelling, (2) prioritization, (3) assessment and (4) synthesis. AHP generates a weight for each evaluation criterion according to the decision maker's pairwise comparisons of the criteria. The higher the weight, the more important the corresponding criterion. Next, for a fixed criterion, AHP assigns a score to each option according to the decision maker's pairwise comparisons of the options based on that criterion. The higher the score, the better the performance of the option with respect to the considered criterion. Finally, AHP combines the criteria weights and the options scores, thus determining a global score for each option, and a consequent ranking. The global score for a given option is a weighted sum of the scores it obtained with respect to all the criteria. Explicitly, AHP will be conducted as follows [45]:

Step 1: Compute the vector of criteria weights. In order to compute the weights for the different criteria, AHP starts creating a pairwise comparison matrix A. The matrix A is a m x m real matrix, where m is the number of evaluation criteria considered. Each entry ajk of the matrix A represents the importance of the j-th criterion relative to the k-th criterion. If ajk > 1, then the j-th criterion is more important than the kth criterion, while if ajk < 1, then the j-th criterion is less important than the kth criterion. If two criteria have the same importance, then the entry ajk is 1. The relative importance between two criteria is measured according to a numerical scale from 1 to 9, as shown in Table 1, where it is assumed that the j-th criterion is equally or more important than the k-th criterion.

Table 1
Table of relative scores

Value of a_{jk}	Interpretation
1	*j* and *k* are equally important
3	*j* is slightly more important than *k*
5	*j* is more important than *k*
7	*j* is strongly more important than *k*
9	*j* is absolutely more important than *k*

Step 2: Compute the matrix of criteria scores. Firstly, the eigenvalue(s) of matrix A needs to be calculated. The scores or weights obtained by the criteria are derived from the eigenvector of matrix A that corresponds to the largest eigenvalue of matrix A.

Step 3: Check the consistency of the pairwise comparison matrix. When many pairwise comparisons are performed, some inconsistencies may typically arise. AHP incorporates an effective technique for checking the consistency of the evaluations made by the decision maker through a consistency index CI where

$$CI = \frac{x - m}{m - 1} \tag{13}$$

A perfectly consistent decision maker should always obtain *CI* = 0, but small values of inconsistency may be tolerated. In particular, if the inconsistencies are tolerable, and a reliable result may be expected from the AHP comparison matrix. RI is the Random Index, i.e. the consistency index when the entries of matrix A are completely random. The values of RI for small problems ($m \leq 10$) are shown in Table 2.

Table 2
Values of the Random Index (RI) for small problems.

m	2	3	4	5	6	7	8	9	10
RI	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.51

One major problem with the AHP process is the consistency of the pairwise comparison matrices [48-50]. To address this problem a proposed pre-Likert scale-AHP approach by Nazri *et al*. [51] manages to solve the issue. In their approach, instead of asking decision makers to directly perform pairwise comparisons among the attributes and criteria, the decision makers will be asked to only rank the level of importance of each attribute or criteria in determining the final decision for any particular problem using Saaty's Likert scale ranging from 1 (least important) to 9 (extremely important). Later, the evaluations from the Likert Scale are converted into Saaty's pairwise comparison table. By doing so, the pairwise comparisons will always be consistent regardless of the number of attributes being analysed. Specifically, Nazri *et al*. [51]'s approach is as follows:

Suppose we have M criteria. Each evaluator must then rate the level of importance of each criterion in determining the weight of that criterion towards the final goal. Suppose that the evaluator rates criterion *i* as w_i and criterion j as w_j. Then a_{ij} which is the pairwise comparison value between criterion i and criterion j will be determined as follows:

$$\begin{aligned} &\text{Let } b = w_i - w_j \\ &\text{If } b > 0 \text{ then } a_{ij} = b+1 \\ &\text{If } b = 0 \text{ then } a_{ij} = 1 \\ &\text{If } b < 0 \text{ then } a_{ij} = 1/(1-b) \end{aligned} \tag{14}$$

These a_{ij}-values will then be transferred into the pairwise matrix A. Once the pairwise matrix A is obtained, the weight for each criterion will be calculated using the existing AHP technique. The process of course includes the consistency test.

7. Conclusion

From the review above, although some of the AI-based techniques are reported to produce superior credit scoring models, in terms of user-friendliness and the simplicity of usage, the techniques are still behind those from the statistical-based techniques. As such, the statistical-based techniques are still the methods of choice for bankers. Among the techniques, the most popular one is the logistic regression. However, the variables used must be carefully selected and the weights given to the variables must be carefully determined. Unfortunately, the model developed can only be verified with the availability of previous historical data. If the data are not available, then one suitable method of choice to determine the relevant criteria and the suitable weights for the criteria will be through the combination of DEMATEL and pre-Likert scale AHP whereby the judgments will be executed by experts who are directly involved in performing this credit screening task. The process of combining DEMATEL and pre-Likert scale AHP is summarized by the flowchart in Appendix A.

References

[1] Samreen, Asia, and Farheen Batul Zaidi. "Design and development of credit scoring model for the commercial banks of Pakistan: Forecasting creditworthiness of individual borrowers." *International Journal of Business and Social Science* 3, no. 17 (2012).

[2] De Bock, Reinout, and Mr Alexander Demyanets. *Bank asset quality in emerging markets: Determinants and spillovers*. No. 12-71. International Monetary Fund, 2012.

[3] Karim, Mohd Zaini Abd, Sok-Gee Chan, and Sallahudin Hassan. "Bank efficiency and non-performing loans: Evidence from Malaysia and Singapore." *Prague Economic Papers* 2, no. 2010 (2010): 118-132.

[4] Vojtek, Martin, and Evžen Koèenda. "Credit-scoring methods." *Czech Journal of Economics and Finance (Finance a uver)* 56, no. 3-4 (2006): 152-167.

[5] Mpofu, T. Peter and M. Mukosera. "Credit scoring techniques: a survey." *International journal of science and research* 3 no. 8 (2012): 165-168.

[6] Bansal, Gaurav, Atish P. Sinha, and Huimin Zhao. "Tuning data mining methods for cost-sensitive regression: a study in loan charge-off forecasting." *Journal of Management Information Systems* 25, no. 3 (2008): 315-336.

[7] Poulsen, J. and A. French. "Discriminant function analysis (DA)." *Journal of forensic sciences* 56 (1996): 297–301.

[8] Koh, Hian Chye, Wei Chin Tan, and Chwee Peng Goh. "A two-step method to construct credit scoring models with data mining techniques." *International Journal of Business and Information* 1, no. 1 (2015).
[9] Mircea, G., M. Pirtea, M. Neamtu, and S. Bazavan. "Discriminant analysis in a credit scoring model." *Paper of Faculty of Economics and Business Administration West University of Timisoara, Romania* (2011).
[10] Abdou, Hussein A., and John Pointon. "Credit scoring, statistical techniques and evaluation criteria: A review of the literature." *Intelligent Systems in Accounting, Finance and Management* 18, no. 2-3 (2011): 59-88.
[11] Bastos, J. (2007). Credit scoring with boosted decision trees, (8156). Retrieved from http://mpra.ub.uni-muenchen.de/8034/\nhttp://mpra.ub.uni-uenchen.de/8034/1/paper.pdf
[12] Zhang, Heping, Richard S. Legro, Jeffrey Zhang, Leon Zhang, Xiang Chen, Hao Huang, Peter R. Casson et al. "Decision trees for identifying predictors of treatment effectiveness in clinical trials and its application to ovulation in a study of women with polycystic ovary syndrome." *Human Reproduction* 25, no. 10 (2010): 2612-2621.
[13] Smith, L. Douglas, and Edward C. Lawrence. "Forecasting losses on a liquidating long-term loan portfolio." *Journal of Banking & Finance* 19, no. 6 (1995): 959-985.
[14] Greenidge, Kevin, and Tiffany Grosvenor. "FORECASTING NON-PERFORMING LOANS IN BARBADOS." *Journal of Business, Finance & Economics in Emerging Economies* 5, no. 1 (2010).
[15] Bolton, Christine. "Logistic regression and its application in credit scoring." PhD diss., 2010.
[16] Anderson, Billie, and J.M. Hardin, Development of Credit Scoring Applications Using SAS® Enterprise Miner™ Course Notes, Cary, NC: SAS Institute Inc., 2009.
[17] Charitou, Andreas, Evi Neophytou, and Chris Charalambous. "Predicting corporate failure: empirical evidence for the UK." *European Accounting Review* 13, no. 3 (2004): 465-497.
[18] Gardner, Mona J., and Dixie L. Mills. "Evaluating the likelihood of default on delinquent loans." *Financial Management* (1989): 55-63.
[19] Angelini, Eliana, Giacomo di Tollo, and Andrea Roli. "A neural network approach for credit risk evaluation." *The quarterly review of economics and finance* 48, no. 4 (2008): 733-755.
[20] Khashman, Adnan. "Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes." *Expert Systems with Applications* 37, no. 9 (2010): 6233-6239.
[21] Alabi, M. A., S. Issa, and R. B. Afolayan. "An application of artificial intelligent neural network and discriminant analyses on credit scoring." *Journal of Modern Mathematics and Statistics* 7, no. 4 (2013): 47-54.
[22] Eletter, Shorouq Fathi, Saad Ghaleb Yaseen, and Ghaleb Awad Elrefae. "Neuro-based artificial intelligence model for loan decisions." *American Journal of Economics and Business Administration* 2, no. 1 (2010): 27-34.
[23] Correa, Alehandro, Andres Gonzalez, and Camilo Ladino. "Genetic Algorithm Optimization for Selecting the Best Architecture of a Multi-Layer Perceptron Neural Network: A Credit Scoring Case." In *SAS Global Forum 2011 Data Mining and Text Analytics*. 2011.
[24] Finlay, S.M. "Using genetic algorithms to develop scoring models for alternative measures of performance." *Lancaster (UK)* (2006).
[25] Finlay, Steven. "Are we modelling the right thing? The impact of incorrect problem specification in credit scoring." *Expert Systems with Applications* 36, no. 5 (2009): 9065-9071.
[26] Kamalloo, Ehsan, and Mohammad Saniee Abadeh. "Credit risk prediction using fuzzy immune learning." *Advances in Fuzzy Systems* 2014 (2014): 7.
[27] Kamalloo, Ehsan, and Mohammad Saniee Abadeh. "Comprehensible credit scoring with fuzzy artificial immune system." In *Electrical Engineering (ICEE), 2010 18th Iranian Conference on*, pp. 542-547. IEEE, 2010.
[28] Shi, Jian, Shu-you Zhang, and Le-miao Qiu. "Credit scoring by feature-weighted support vector machines." *Journal of Zhejiang University SCIENCE C* 14, no. 3 (2013): 197-204.
[29] Bellotti, Tony, and Jonathan Crook. "Support vector machines for credit scoring and discovery of significant features." *Expert Systems with Applications* 36, no. 2 (2009): 3302-3308.
[30] Shieh, Jiunn-I., Hsin-Hung Wu, and Kuan-Kai Huang. "A DEMATEL method in identifying key success factors of hospital service quality." *Knowledge-Based Systems* 23, no. 3 (2010): 277-282.
[31] Tzeng, Gwo-Hshiung, Cheng-Hsin Chiang, and Chung-Wei Li. "Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL." *Expert systems with Applications* 32, no. 4 (2007): 1028-1044.
[32] Wu, Hsin-Hung, and Shih-Yu Chang. "A case study of using DEMATEL method to identify critical factors in green supply chain management." *Applied Mathematics and Computation* 256 (2015): 394-403.
[33] Siddiqi, Naeem. *Credit risk scorecards: developing and implementing intelligent credit scoring*. Vol. 3. John Wiley & Sons, 2012.
[34] Edwards, S. "In too deep: CAB clients' experience of debt (Citizens Advice)." (2003).
[35] Gärling, T., E. Kirchler, A. Lewis, and F. V. Raaij, F. Van. (2010). "Psychology , financial decision making , and financial crises." Unpublished manuscript, University of Gothenburg, (2010), 1–81.

[36] McHugh, Sandie, and Rob Ranyard. "Credit repayment decisions: The role of long-term consequence information, economic and psychological factors." *Review of Behavioural Finance* 4, no. 2 (2012): 98-112.
[37] Valins, Oliver. *When debt becomes a problem: A literature Study*. Ministry of Social Development, 2004.
[38] Wang, Lili, Wei Lv, and Lin Jiang. "The impact of attitude variables on the credit debt behavior." *Nankai Business Review International* 2, no. 2 (2011): 120-139.
[39] Dearden, Chris, Jackie Goode, Grahame Whitfield, and Lynne Cox. "Credit and debt in low-income families." (2010).
[40] Turner, Michael, Alyssa Stewart Lee, Ann Schnare, Robin Varghese, and Patrick D. Walker. "Give credit where credit is due: Increasing access to affordable mainstream credit using alternative data." *Chapel Hill, NC and Washington, DC: Political and Economic Research Council and The Brookings Institution* (2006).
[41] Norgren, Claes. "The causes of the global financial crisis and their implications for supreme audit institutions." *Auditor General of the Swedish National Audit Office Stockholm* (2010).
[42] Leonard, Kevin J. "The development of credit scoring quality measures for consumer credit applications." *International Journal of Quality & Reliability Management* 12, no. 4 (1995): 79-85.
[43] Siddiqi, Naeem. *Credit risk scorecards: developing and implementing intelligent credit scoring*. Vol. 3. John Wiley & Sons, 2012.
[44] Hartati, Sri, Retantyo Wardoyo, Agus Harjoko, Jl Palembang-prabumulih, and Ogan Ilir. "ELECTRE methods in solving group decision support system bioinformatics on gene mutation detection simulation." (2011).
[45] Saaty, T.L. “The analytic hierarchy process.” McGraw-Hill, New York, 1980.
[46] Bhattarai, Shashi, and Shivjee Roy Yadav. "AHP application in banking: unfolding utility in a situation of financial crisis." In *Proceedings of the international symposium on the analytic hierarchy process, ISAHP2013*. 2009.
[47] Gutiérrez-Nieto, Begoña, Carlos Serrano-Cinca, and Juan Camón-Cala. "A credit score system for socially responsible lending." *Journal of Business Ethics* 133, no. 4 (2016): 691-701.
[48] Noble, Elizabeth E., and Pedro P. Sanchez. "A note on the information content of a consistent pairwise comparison judgment matrix of an AHP decision maker." *Theory and Decision* 34, no. 2 (1993): 99-108.
[49] Tummala, VM Rao, and Yat-wah Wan. "On the mean random inconsistency index of analytic hierarchy process (AHP)." *Computers & industrial engineering* 27, no. 1-4 (1994): 401-404.
[50] Ozdemir, Müjgan Sagir. "Validity and inconsistency in the analytic hierarchy process." *Applied Mathematics and Computation* 161, no. 3 (2005): 707-720.
[51] Nazri, E. M., M. Balhuwaisl, and M. M. Kasim. “A pre-evaluation step towards a guaranteed consistent AHP-based pairwise comparison.” *Journal of Advanced Research in Social and Behavioural Sciences* 4 no. 1, (2016): 73-80.
[52] Greene, William. "Sample selection in credit-scoring models." *Japan and the world Economy* 10, no. 3 (1998): 299-316.
[53] Witzany, Jiří. *Credit risk management and modeling*. Praha: Oeconomica, 2010.
[54] Choy, Murphy, and Ma Nang Laik. "A Markov Chain approach to determine the optimal performance period and bad definition for credit scorecard." *International Journal’s Research Journal of Social Science and Management* 1 (2011): 227-234.
[55] Gurný, Petr, and Martin Gurný. "Comparison of Credit Scoring models on probability of default estimation for US Banks." *Prague economic papers* 2 (2013): 163-181.
[56] Sumrit, Detcharat, and Pongpun Anuntavoranich. "Using DEMATEL method to analyze the causal relations on technological innovation capability evaluation factors in Thai technology-based firms." *Int Trans J Eng Manag Appl Sci Technol* 4, no. 2 (2013): 081-103.
[57] Lee, Hakyeon, Sora Lee, and Yongtae Park. "Selection of technology acquisition mode using the analytic network process." *Mathematical and Computer Modelling* 49, no. 5 (2009): 1274-1282.
[58] Yang, Yu-Ping Ou, How-Ming Shieh, Jun-Der Leu, and Gwo-Hshiung Tzeng. "A novel hybrid MCDM model combined with DEMATEL and ANP with applications." *International journal of operations research* 5, no. 3 (2008): 160-168.
[59] Fausett, Laurene V. *Fundamentals of neural networks*. Prentice-Hall, 1994.
[60] Back, Barbro, Teija Laitinen, and Kaisa Sere. "Neural networks and genetic algorithms for bankruptcy predictions." *Expert Systems with Applications* 11, no. 4 (1996): 407-413.
[61] Leung, Kevin, France Cheong, and Christopher Cheong. "Consumer credit scoring using an artificial immune system algorithm." In *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*, pp. 3377-3384. IEEE, 2007.