

The implication of machine learning for financial solvency prediction: an empirical analysis on public listed companies of Bangladesh

Financial
solvency
prediction

303

Mohammad Abdullah
*Faculty of Business and Management, Universiti Sultan Zainal Abidin,
Kuala Terengganu, Malaysia*

Received 29 November 2020
Revised 31 January 2021
24 March 2021
26 April 2021
2 June 2021
Accepted 2 June 2021

Abstract

Purpose – Financial health of a corporation is a great concern for every investor level and decision-makers. For many years, financial solvency prediction is a significant issue throughout academia, precisely in finance. This requirement leads this study to check whether machine learning can be implemented in financial solvency prediction.

Design/methodology/approach – This study analyzed 244 Dhaka stock exchange public-listed companies over the 2015–2019 period, and two subsets of data are also developed as training and testing datasets. For machine learning model building, samples are classified as secure, healthy and insolvent by the Altman Z-score. R statistical software is used to make predictive models of five classifiers and all model performances are measured with different performance metrics such as logarithmic loss (logLoss), area under the curve (AUC), precision recall AUC (prAUC), accuracy, kappa, sensitivity and specificity.

Findings – This study found that the artificial neural network classifier has 88% accuracy and sensitivity rate; also, AUC for this model is 96%. However, the ensemble classifier outperforms all other models by considering logLoss and other metrics.

Research limitations/implications – The major result of this study can be implicated to the financial institution for credit scoring, credit rating and loan classification, etc. And other companies can implement machine learning models to their enterprise resource planning software to trace their financial solvency.

Practical implications – Finally, a predictive application is developed through training a model with 1,200 observations and making it available for all rational and novice investors (Abdullah, 2020).

Originality/value – This study found that, with the best of author expertise, the author did not find any studies regarding machine learning research of financial solvency that examines a comparable number of a dataset, with all these models in Bangladesh.

Keywords Financial distress, Machine learning, Artificial neural network, Ensemble classifier, Bankruptcy prediction

Paper type Research paper

JEL Classification — G30, G17, G33

© Mohammad Abdullah. Published in *Journal of Asian Business and Economic Studies*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

The author would like to express his very great appreciation to Dr. Bryan Weber (Assistant Professor, College of Staten Island – CUNY, Department of Economics) for his valuable and constructive copy editing of this paper. His willingness to give his time so generously has been very much appreciated.



1. Introduction

Nowadays, the probability and number of insolvent firms have been significantly increasing. As a result, financial solvency prediction is a significant concern throughout academia, primarily in finance. Subsequently, a firm's solvency is extremely important for investors, creditors, stockholders, insurance policyholders, tender suppliers, investment managers, financiers, governments and capital market investors, etc. Thus, a well-developed guideline to effectively assess the probability of financial insolvency is highly desired. Over the last several decades, two models have been developed for predicting corporate failure using discriminant analysis and logistic regression (Altman, 1968; Ohlson, 1980). The forecast of financial distress is a vital and interesting agenda that has been served as the inspiration for numerous academic studies over the last five decades. Due to the high impact of corporate failure events, researchers in artificial intelligence urgently apply intelligent methods to forecast bankruptcy.

Machine learning algorithms evaluate models by training and predicting the output of using previous experience. In this study, machine learning models are trained to distinguish solvent and insolvent firms based on different characteristics of the business including, growth, profitability, leverage, market value, liquidity and growth measures. The corporation becomes insolvent due to different quantifiable and nonquantifiable factors. Using quantifiable factors, many researchers have been trying to predict financial distress for many decades. From the previous literature, the author found that ratio analysis is a significant predictor of financial solvency (Altman, 1968; Beaver, 1966).

To the knowledge of the author, after analyzing the previous literature, there is no application in Bangladesh regarding machine learning for financial solvency prediction. Bangladesh has shifted from an agriculture-based economy to an industrialized economy. In addition, the country's irrational investors are becoming rational with the sound knowledge of investment. Using this background, machine learning algorithms are applied to real-world problems, for example, (1) lender organizations that can leverage their investment decisions, (2) policymakers that could better recognize and examine strategic investment decisions, (3) individual-level investor wants to buy insurance policy form a solvent company and (4) advance payment of goods for tender in a solvent company; based on the outcomes of machine learning models all of these decision-makers will be able to make a decision confidently. Therefore, we have to construct a systematic solution for financial solvency prediction. This research aims to "*predict financial solvency by using various machine learning models and linking them together to compare their accuracy and develop a simple model for rational investment decision-makers for Bangladesh*".

From the trained machine learning models, these models will be easily replicated for future prediction by inputting their raw data, not only by scholars but also by finance professionals and rational investors. This study is a modest attempt to implement machine learning for predicting financial solvency, and it also develops a simple approach to make it more user-friendly for all levels of rational investment decision-makers in Bangladesh. The objective of this study is to develop machine learning models to predict the financial solvency of businesses for rational investment decision-makers. The evaluation of different algorithms makes it available to the end user and can be considered as the main contributions of this study. The outcome of this research will be beneficial to the industry practitioners as well as rational investors in their decision-making process

2. Theoretical background

2.1 Machine learning and financial solvency prediction

The early studies regarding solvency prediction were conducted within different streams of research. Those studies concentrated on a specific set of ratios and occasionally contrasted

ratios of failed corporations with those of successful corporations. Altman (1968), Beaver (1966) and Ohlson (1980) have developed the foundation of bankruptcy prediction models, which have been used for prospective financial solvency prediction and model development. Machine learning enables processors to find gist data from big data automatically (Michie *et al.*, 1994; Shapiro, 2001). A recent bibliometric analysis of machine learning models for intelligent bankruptcy prediction methods for corporate firms finds that the first machine learning models for predicting bankruptcy were placed in 1991 according to Web of Science. Though the number of publications gradually increased in 2008, considerable growth in publication counts can be seen from 2008 to 2009. This emphasizes the growth of this topic from the financial crisis in 2007–2008 period, and it shows the great demand for machine learning techniques and their applications in financial distress prediction (Shi and Li, 2009). The author also found that cooperation between scholars is weak, particularly at the global level. As a result, the majority of publications are from the USA, China, Taiwan, Spain and South Korea. The concept of machine learning to predict financial solvency has been previously studied by Tam using neural network models (Tam, 1991). Nevertheless, Jabeur *et al.* (2020) found that machine learning models can be used more effectively for bond rating than conventional methods by analyzing classifiers (Jabeur *et al.*, 2020). Table 1 demonstrates a list of different studies that emphasized bankruptcy prediction, and they are summarized by each study based on the sample size, time, models used and their significant findings.

2.2 Financial solvency prediction and Altman Z Score

Altman (1968) developed a multivariate discriminant analysis model for financial distress prediction also known as the Z-Score model. After four decades with many arguments and adjustments, recent models are demonstrated as in models 1 and 2, while model 1 is for nonmanufacturing firms and model 2 is used for other companies. x_1 = working capital/total assets, x_2 = retained earnings/total assets, x_3 = earnings before interest and taxes/total assets, x_4 = market value of equity/total liabilities, x_5 = total sales/total assets.

$$Z = 6.56x_1 + 3.26x_2 + 6.72x_3 + 1.05x_4 \quad (1)$$

$$Z' = 1.26x_1 + 1.4x_2 + 3.3x_3 + 0.6x_4 + 1.0x_5 \quad (2)$$

The function used in the Z-score is demonstrated in equations (1) and (2), where equation (2) is for the value of Z-Score companies that can be categorized into a different category as the red zone, safe zone, and gray zone (Altman, 1968). In this study, the Altman Z-score will be used for company classification.

3. Data collection and methodology

This study is based on the context of Bangladesh. There are 323 companies in the Dhaka Stock Exchange whose stocks are traded daily. This study tries to cover all public listed companies in the Dhaka stock exchange in the dataset. However, a few companies' information is not available. As a result, 244 companies have been selected as samples. This study used big data mining techniques to analyze a large dataset of annual reports of Bangladeshi companies through mining data set from 2015 to 2019 from the annual report and Dhaka Stock Exchange reports.

For data analysis, all 244 companies' data in the 2015–2019 period are classified into three categories by implementing the Altman Z-score. Equation (1) is applied for manufacturing-related business such as textile, pharmaceuticals, engineering, tannery, paper and printing, food, ceramics sector, etc. Then, equation (2) is used for service industries such as banks, financial institutions, insurance, telecommunication, etc. Classification criteria from Table 1 are used for classification (Altman, 1968).

Study	Sample size	Period	Models ***	Findings
Beaver (1966)	275	1949–1963	FR	Accounting data can be utilized for prediction
Altman (1968)	66	1946–1965	MDA	MDA model can predict corporate failure correctly with 94% accuracy
Ohlson (1980)	105	1970–1976	LR	Significant improvement may require additional predictors
Raghupathi <i>et al.</i> (1991)	129	1975–1982	ANN	NN has a significant role in predicting failure
Sun and Shenoy (2007)	7,822	1989–2002	BN	Best when the number of states for discretization is 2 or 3
Chandra <i>et al.</i> (2009)	16,816	1962–1999	MLP, CART, LR, RF, SVM, BS	<i>t</i> -statistic used for feature selection and ten features are mined
Ioannidis <i>et al.</i> (2010)	944	2007–2008	MLP, CART, KNN, LR, SM	Multi-criteria decision aid and artificial neural networks (ANNs) achieve the highest accuracies
Cecchini <i>et al.</i> (2010)	156	1994–1999	SVM	Bankruptcy (83.87%) and fraud (81.97%) with the combined data
Kim and Kang (2010)	1,458	2002–2005	SVM, Adaboost, GMBost	Bagged and the boosted neural networks showed improved performance over traditional neural networks
Yeh <i>et al.</i> (2010)	114	2005–2007	DEA + Rough sets + SVM	DEA provides valuable information in predictions
Chen (2011)	100	2000–2007	PCA, DT, LR	The artificial intelligent approach could be a more suitable methodology than traditional statistics
Olson <i>et al.</i> (2012)	1,321	2005–2009	DT, LR, MLP, RBFN, SVM	Decision trees were relatively more accurate compared to neural networks and support vector machines
du Jardin and Séverin (2012)	17,480	1998–2004	MDA, LR, ANN	Kohonen map can be used as a prediction model
Booth <i>et al.</i> (2014)		2000–2012	LR, DT, ANN, SVR, RF	Recency-weighted ensembles of random forests produce better results
Heo (2014)	29,862	2012–2018	AdaBoost, ANN, SVM, DT	The AdaBoost has more analytical power than others
Geng <i>et al.</i> (2015)	214	2001–2008	ANN, DT, SVM, MDA	Neural networks provide the highest accuracy and are robust to experimental conditions
Danenas and Garsva (2015)	21,487	2000–2007	SVM, ANN, LR	SVM technique is capable to produce results, comparable to other classifiers, such as logistic regression and RBF network
Liang <i>et al.</i> (2016)	478	1999–2009	SVM, KNN, NB, CART, MLP	Classifications of the board and ownership composition are the most important features
Xiao <i>et al.</i> (2016)	1,000		SVM, BG, DT, LR	Weighted voting generates diverse and locally accurate base classifiers
Barboza <i>et al.</i> (2017)	13,300	1985–2013	SVM, BG, BO, RF, ANN, MDA, LR	ML models can be utilized for creating a model with better classification correctness
Jones (2017)	1,115	1987–2013	GMBost	CEO compensation and ownership structure are the strongest predictors

Table 1.
Summary of relevant papers

(continued)

Table 1.

Study	Sample size	Period	Models ***	Findings
Mai et al. (2019)	11,827	1994–2014	ANN, LR, SVM, RF	deep learning can efficiently integrate the incremental information from textual data with numeric information and achieve better prediction accuracy
Hosaka (2019)	2,168	2012–2016	ANN, DT, LR, SVM, Adaboost, MDA	ANN performance is higher compared to other methods

Note(s): Artificial neural networks (ANNs), Bayesian network (BN), classification and regression trees (CART), data envelopment analysis (DEA), decision trees (DT), financial ratios (FR), *K*-nearest neighbor (KNN), logistic regression (LR), multilayer perceptron (MLP), multivariate discriminant analysis (MDA), radial basis function network (RBFN), random forest (RF), stacked models (SM), support vector machines (SVM)

After the multivariate discriminant analysis, all companies in the final sample are categorized as secure, healthy and insolvent based on the *Z*-score. The discrimination zones are explained as follows: (1) if a company is “secure” then there is low risk; (2) the company is solvent if it is categorized as “healthy” which might head to insolvency and (3) if a company is categorized as “insolvent”, it denotes that the company’s financial condition is in the distress zone. After categorizing the company’s solvency, two subsets of data covering the 2016–2017 period are extracted to develop the training sample. Another subset of data is then created as testing data set for the 2018–2019 period. *R* Statistical Computing Software is used for data mining, data analysis and machine learning model training. Five machine learning algorithms are applied for the training and testing of models; they are demonstrated in [Table 3](#).

For the development of models, 11 features have been extracted, and [Table 4](#) presents a demonstration of the features. *X*₁ to *X*₅ is used for the discriminant zone classification and six more features have been extracted for creating features in machine learning models. Previous studies indicate that all of these features have significant relationships and these features are also called predictor variables, and they can explain financial solvency ([Barboza et al., 2017](#); [Mai et al., 2019](#); [Son et al., 2019](#); [Kim et al., 2016](#)). There are many performance indicators for machine learning models and hyperparameter tuning methods. In this study prediction performance will be measured by logarithmic loss (logLoss), area under the curve (AUC), precision recall AUC (prAUC), accuracy, kappa, sensitivity and specificity.

$$\text{Sensitivity} = \frac{\text{NTP}}{\text{NTP} + \text{NFN}} \quad (3)$$

$$\text{Specificity} = \frac{\text{NTN}}{\text{NTN} + \text{NFN}} \quad (4)$$

In [equation \(3\)](#) NTP is the number of true positive classification, that is insolvent firms classified correctly. And NTN is the number of true negative classification, that is secure firms classified correctly. In addition, in [equation \(4\)](#), NFN is the number of false-positive classification, which insolvent firms classified incorrectly, while NFP is the number of false positive classification, which denotes secure firms classified incorrectly. The value of sensitivity and specificity will be close to 1 if there is low classification error. After the model training and prediction, all performance metrics will be analyzed and compared between all models’ results.

4. Results and discussion

To examine the prediction of solvency by machine learning models, it is necessary to classify the samples into three classes, which are defined in [Table 2](#). After the calculation of the

Z-score value, classes are defined as per the discriminant zones. The industry-wide data over the 2015–2019 period are demonstrated in Figure 1. There are a smaller number of “secure” companies out of the sample over the given period. Also, nonbank financial institutions are facing major insolvency problems as most of them are categorized as “insolvent”. Most interestingly, almost every insurance company performs well. Also, the textile industry is in a good position, but its performance worsens yearly. A previously conducted study found that only Sharia-based banks are performing better than conventional banks, and the result are consistent with our findings (Abdullah, 2015).

The next step is descriptive statistics. In Table 5, descriptive statistics of three datasets are represented, and all values are rounded to two decimals. The data are collected from 244 companies of 18 different industries. This study measures the performance of different models for solvency prediction; no data transformation and no normalization were conducted (Barboza *et al.*, 2017). The current ratio (CR) has highly deviated, and X4 is highly deviated, which indicates that the market values of companies are much higher than their liabilities. The skewness and kurtosis indicate that data are highly skewed. As industry-wise companies’ market value, profitability and company size are distinct; this will be a better set for training and testing models because this output is a better prediction for all companies. Also, this full dataset can be used for future model training and prediction for all Bangladeshi companies.

Table 2.
Discriminant zones

Classes	Manufacturing industry	Nonmanufacturing industry
Secure	$Z > 2.90$	$Z > 2.60$
Healthy	$1.23 < Z < 2.90$	$1.10 < Z < 2.60$
Insolvent	$Z < 1.23$	$Z < 1.10$

Table 3.
Machine learning
models for this study

Machine learning algorithm	Base learners
Artificial neural network classifier (ANNC)	Averaged neural network
Support vector machines (SVM)	Support vector machines with linear kernel
Naive Bayes classifier (NBC)	Naive Bayes
<i>K</i> -nearest neighbor classifier (KNNC)	<i>K</i> -nearest neighbors
Ensemble classifier (EC)	Bagged CART

Table 4.
Selected variables/
features for predictive
modeling

Variables	Description
X1	Working capital/total assets
X2	Retained earnings/total assets
X3	Earnings before interest and taxes/total assets
X4	Market value of equity/total liabilities
X5	Total sales/total assets
CR	Current assets/current liabilities
CLTL	Current liabilities/total liabilities
NIS	Earnings after tax/total sales
TETA	Total equity/total assets
ROA	Earnings after tax/total assets
ROE	Earnings after tax/total equity

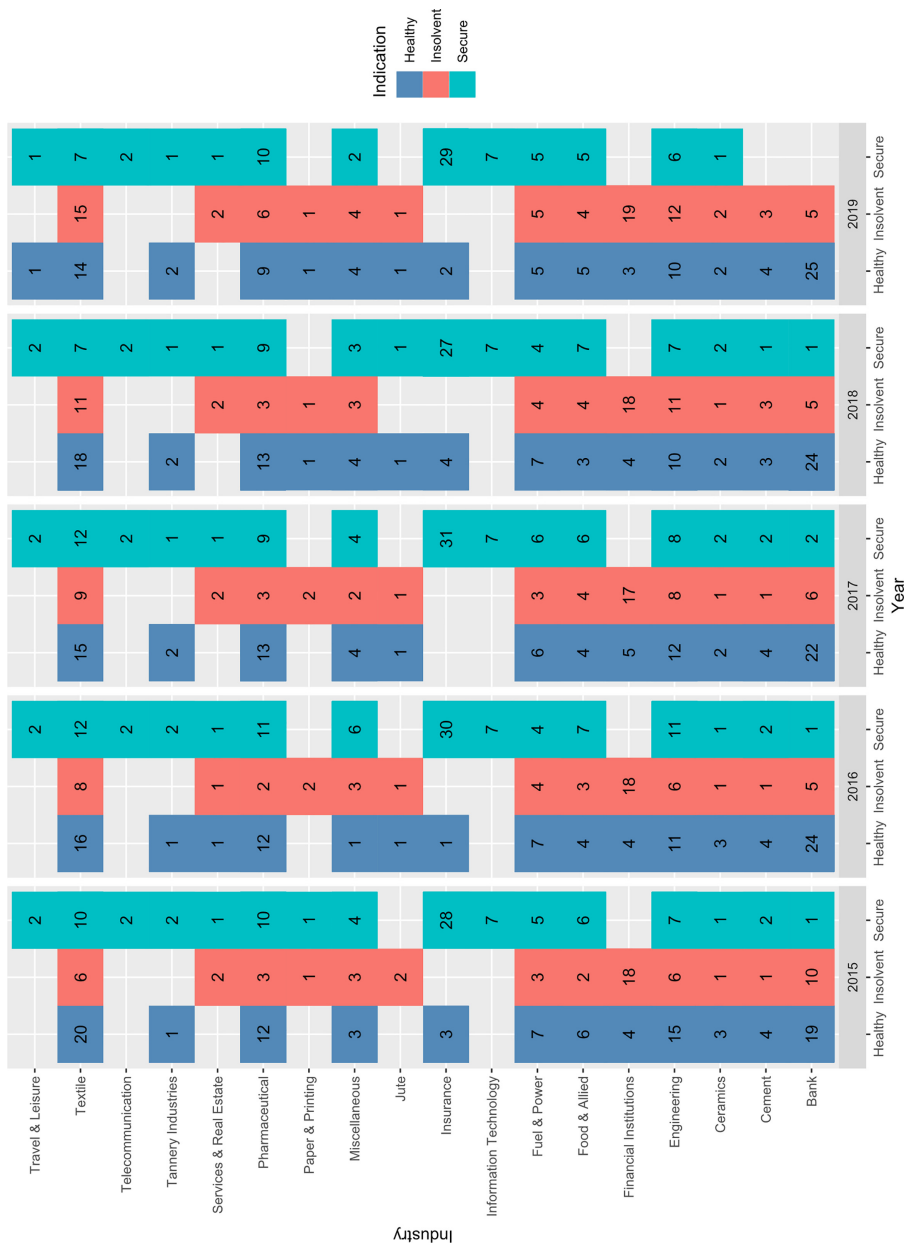


Figure 1. Yearly classification by industry

It is vital to check the correlation between all variables of every dataset before moving to predictive machine learning model training. Pearson correlation has been analyzed to check the correlation between all predictor variables. The correlation matrix plot of all datasets is plotted in the heatmap of Figure 2, and results are presented in Table 6. From the full dataset correlation, results indicate almost all correlations are significant except the CR. The CR has a significant correlation with current liabilities/total liabilities (CLTL) and X1 only. Moreover, the results of the training dataset correlation analysis are presented in Table 6. The results indicate that X1 has a significant correlation with most of the variables except X2, X5, NIS and ROE. However, X2 has a significant correlation with most of the variables. Results are similar to the full dataset and only the CR has less correlation. Therefore, the training sample is consistent with full data.

Finally, the results of the correlation analysis of the testing dataset are demonstrated in Table 6. The results of the analysis indicate the same results of the full dataset and training dataset. Most of all variable's correlations are significant. In summary, most of the predictor variables are significantly correlated, and the dataset can be used for building machine learning models.

4.1 Artificial neural network classifier results

For ANNC training, averaged neural network base learner is used for financial solvency prediction. In this research, ten-fold classification and five neurons are used for model training. The confusion matrix of this trained model is demonstrated in Figure 3. In the

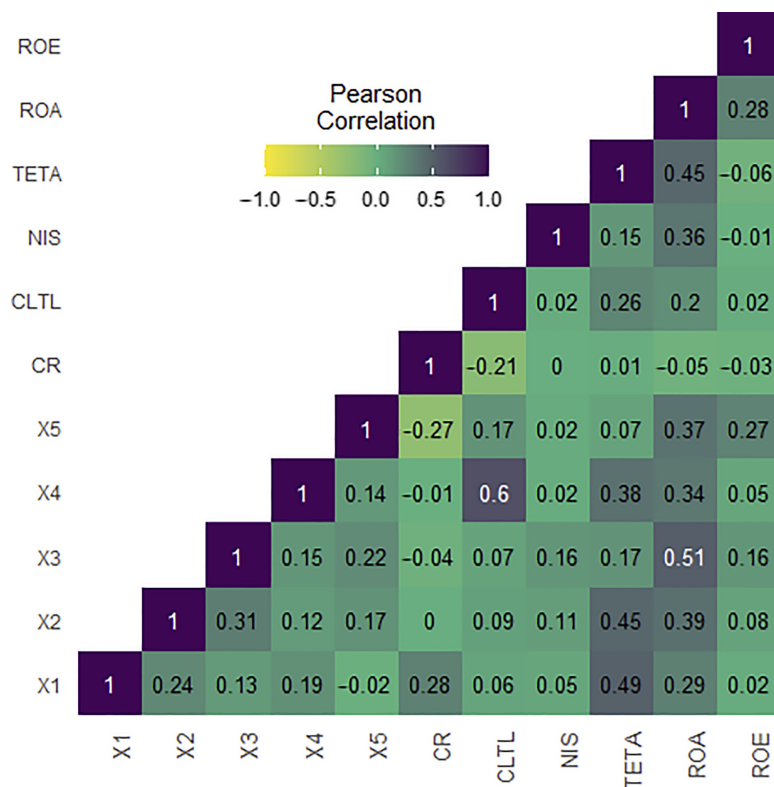


Figure 2. Correlation matrix of full dataset

Table 6.
Correlations between
all predictor variables
of all datasets

	X1	X2	X3	X4	X5	CR	CLTL	NIS	TETA	ROA	ROE
<i>Correlations of Full Dataset</i>											
X1	1										
X2	0.236**	1									
X3	0.129**	0.309**	1								
X4	0.191**	0.118**	0.153**	1							
X5	-0.025	0.167**	0.224**	0.144**	1						
CR	-0.002	-0.002	-0.035	-0.006	-0.265**	1					
CLTL	0.063*	0.091**	0.072**	0.598**	0.175**	-0.208**	1				
NIS	0.047	0.112**	0.162**	0.023	0.017	0.001	0.019	1			
TETA	0.490**	0.450**	0.172**	0.381**	0.068*	0.008	0.263**	0.152**	1		
ROA	0.293**	0.394**	0.512**	0.337**	0.372**	-0.047	0.202**	0.361**	0.445**	1	
ROE	0.023	0.085**	0.161**	0.047	0.269**	-0.031	0.023	-0.007	-0.063*	0.277**	1
<i>Correlations of Training Dataset</i>											
X1	1										
X2	0.043	1									
X3	0.175	0.217**	1								
X4	0.199**	0.091*	0.238**	1							
X5	-0.063	0.148**	0.453**	0.145**	1						
CR	0.294	-0.008	-0.108*	-0.039	-0.292**	1					
CLTL	0.096*	0.074	0.117**	0.702**	0.156**	-0.180**	1				
NIS	0.046	0.112**	0.433**	0.023	0.024	-0.029	0.024	1			
TETA	0.354	0.348**	0.297**	0.418**	0.076	-0.040	0.281**	0.198**	1		
ROA	0.202**	0.303**	0.872**	0.319**	0.397**	-0.077	0.182**	0.550**	0.394**	1	
ROE	-0.015	0.067	0.262**	0.041	0.380**	-0.063	0.021	-0.138**	-0.161**	0.259**	1
<i>Correlations of Testing Dataset</i>											
X1	1										
X2	0.504**	1									
X3	0.105*	0.469**	1								
X4	0.165**	0.136**	0.103*	1							
X5	-0.010	0.186**	0.105*	0.133**	1						
CR	0.249**	0.016	-0.001	0.048	-0.250**	1					
CLTL	-0.033	0.170**	0.028	0.278**	0.385**	-0.482**	1				
NIS	0.108*	0.227**	0.071	0.029	-0.010	0.029	0.019	1			
TETA	0.594**	0.648**	0.119**	0.330**	0.059	0.031	0.356**	0.280**	1		
ROA	0.381**	0.569**	0.355**	0.301**	0.324**	-0.014	0.217**	0.265**	0.507**	1	
ROE	0.062	0.106*	0.125**	0.058	0.195**	-0.002	0.020	0.115*	0.007	0.324**	1

Note(s): ** Correlation is significant at the 0.01 level and * Correlation is significant at the 0.05 level

middle of each tile, the normalized count of classes is plotted and the count column percentages are given below. Within the testing dataset, 89.3% was predicted to be secure class, 0.6% insolvent and 10% healthy. At the right portion of each tile, the row percentage of

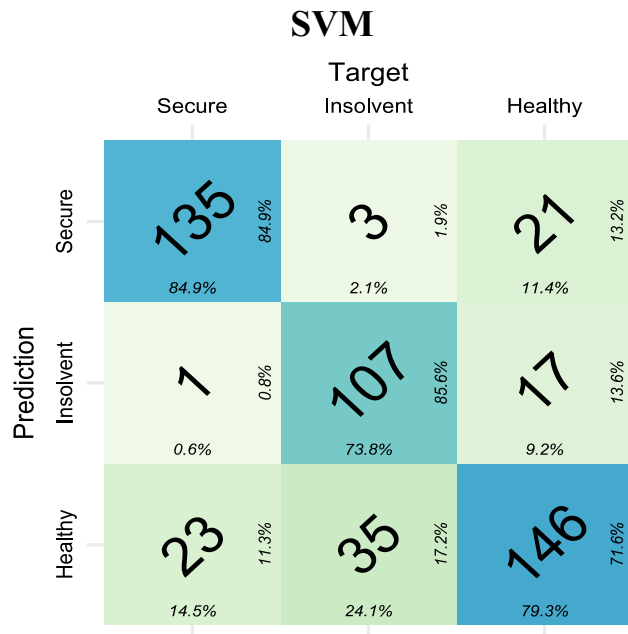
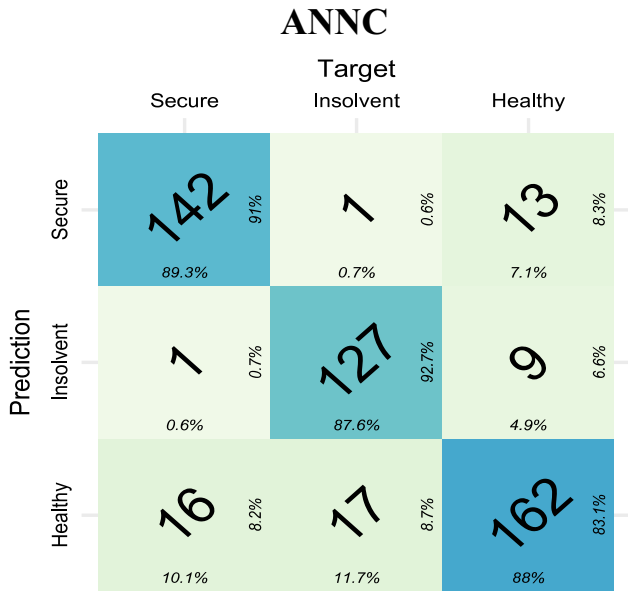


Figure 3. Confusion matrix of ANNC and SVM

all observations is plotted, 91% was “secure”, while 0.7% was “insolvent” and 8.2% “healthy”. The overall results of the confusion matrix of the ANNC model have significant predictive power. Nevertheless, the results of model training and testing are demonstrated in Table 7 as ANNC. The accuracy of this model at the stage of training and testing are 94 and 82%, respectively. In addition, the AUC of training and testing of models are 96 and 94%, respectively. The sensitivity and specificity of training are 84 and 92%, also testing stage results are 88 and 94%. Lastly, prAUC, logloss and kappa have significant results in both training and testing models, and previous studies support the results (Raghupathi *et al.*, 1991; Geng *et al.*, 2015; Barboza *et al.*, 2017; Hosaka, 2019).

4.2 Support vector machine results

In the process of support vector machines model training, a support vector machine with a linear kernel is used as a base learner, and the results summary is shown in Table 7. The accuracy and area under the curve of the SVM model are 94 and 84% for model training, 96 and 88% for model testing. However, the value of sensitivity and specificity for model training is 84 and 92%, 88%, and 94%, respectively for model testing. The confusion matrix of the SVM model is demonstrated in Figure 3; at the bottom of every tile the column percentage is shown, where, for example, 84.9% is predicted as the proportion classified as secure. The right side of every column indicates the value of row percentage. The previous studies and these results indicate that support vector machines can predict financial solvency significantly (Chandra *et al.*, 2009; Cecchini *et al.*, 2010; Kim and Kang, 2010; Xiao *et al.*, 2016).

4.3 Naive Bayes classifier results

For training, the naive Bayes classifier and naive Bayes base learner are used with ten-fold validation. The performance metrics are presented in Table 7. The accuracy of NBC models for training and testing is 60 and 74%, ACU’s percentage is 86 and 89% for training and testing of the model. Also, the sensitivity and specificity for model training are 55 and 79%, 85%, and 92% for model testing. The confusion matrix of the NBC model is plotted in Figure 4. All these results indicate that the naive Bayes classifier can predict the financial solvency of corporations and the previous literature also supports the results (Sun and Shenoy, 2007; Tavana *et al.*, 2013; Masmoudi *et al.*, 2019; Barboza *et al.*, 2017).

4.4 K-nearest neighbor classifier results

K-nearest neighbors is used as a base learner with a ten-fold classification for model training. The confusion matrix of KNNC is plotted in Figure 4, and results of training and testing

Model	logLoss	AUC	prAUC	Accuracy	Kappa	Sensitivity	Specificity
<i>Results of the training dataset</i>							
ANNC	0.56	0.94	0.82	0.84	0.75	0.84	0.92
SVM	0.51	0.93	0.81	0.80	0.69	0.79	0.89
NBC	2.46	0.86	0.66	0.60	0.36	0.55	0.79
KNNC	5.75	0.87	0.13	0.82	0.73	0.82	0.91
EC	18.30	0.98	0.68	0.89	0.84	0.89	0.95
<i>Results of the testing dataset</i>							
ANNC	9.65	0.96	0.94	0.88	0.82	0.88	0.94
SVM	4.71	0.93	0.86	0.80	0.69	0.79	0.89
NBC	8.44	0.89	0.80	0.74	0.61	0.74	0.87
KNNC	22.93	0.88	0.09	0.85	0.77	0.85	0.92
EC	0.43	0.96	0.71	0.86	0.82	0.86	0.93

Table 7. Results of training and prediction of all machine learning models

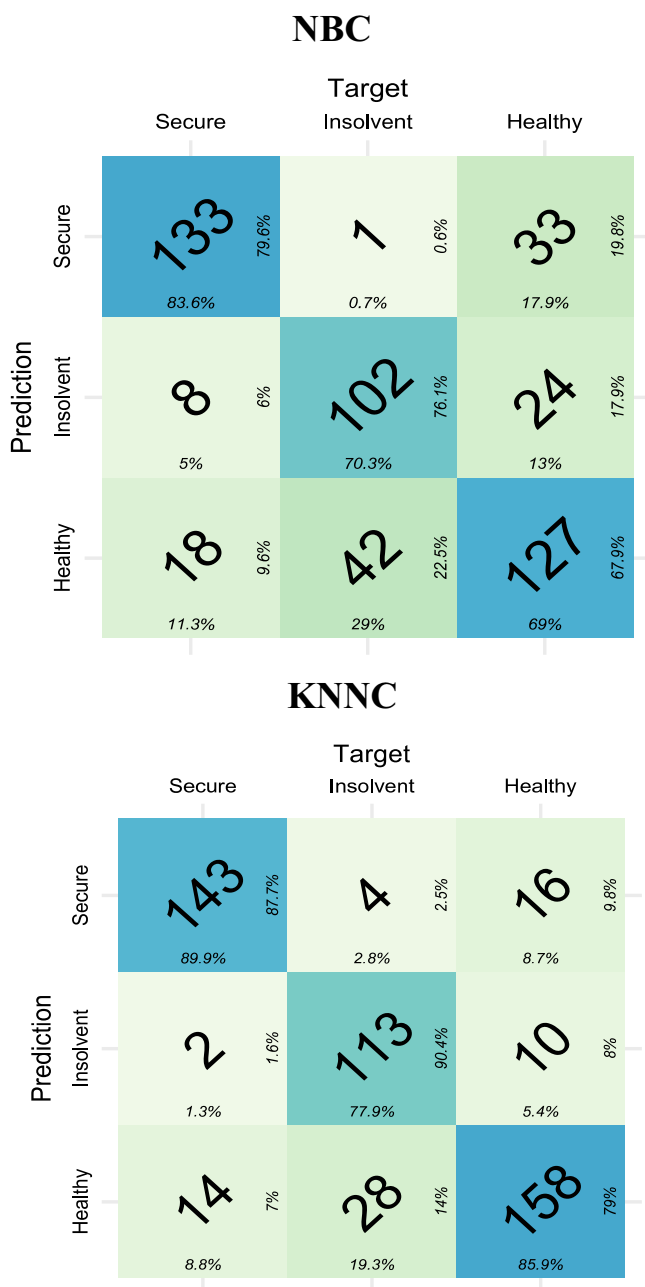


Figure 4. Confusion matrix of NBC and KNNC

models are demonstrated in Table 7. The accuracy of the KNNC model training and testing is 82 and 85%, also AUC values are 87 and 88%. The sensitivity and specificity of model training are 82 and 91%, 85%, and 92% for model testing. Sum of all results indicates that the

4.5 Ensemble classifier results

The final model of this study is Ensemble Classifier, Bagged classification, and regression trees (CART) which is used as a base learner with a ten-fold classification for this model. The confusion matrix of EC is plotted in Figure 5. The results of training and testing of EC models are presented in Table 7, also the confusion matrix is presented in Figure 5. The sensitivity and specificity for model training are 89 and 95% and for testing are 86 and 93%. AUC and accuracy for the training model are 98 and 89% and for testing models are 96 and 86%. Several previous studies also found these consistent results (Heo, 2014; Kim and Kang, 2010; Pisula, 2020). Consequently, this indicates that the ensemble classifier can be used for financial solvency prediction.

In Table 7 all models' performance metrics are presented. logLoss is the measurement for the performance of a machine learning classification model where the prediction class can be a binary class or multiclass. The main goal of a machine learning model is to minimize this logLoss value. When the logLoss is near 0 then the model can be classified as a perfect model. AUC represents the results of the area under the receiver operating characteristic (ROC) curve. The graphical presentation of a classifier with its discriminant threshold value is called the ROC curve; in Figure 6 all models' ROC curves are illustrated. The higher the value of AUC, the higher the predictive power of the model will be. Nevertheless, prAUC calculates the percentage of predicted positive class.

Cohen's kappa statistics are a better indicator for multi-class predictive modeling, which indicates the degree of model fitness of good. According to Landis and Koch (1977) value of kappa can be categorized in a different segment, where the value is less than 0%. It indicates that the model is not a good fit, while 0%–20% is a low and 21%–40% is a fair significance. In addition, 41%–60% indicates moderate significance, while 61%–80% indicates considerable significance, and 81%–100% indicates perfect significance. Sensitivity is the measurement of



Figure 5. Confusion matrix of EC

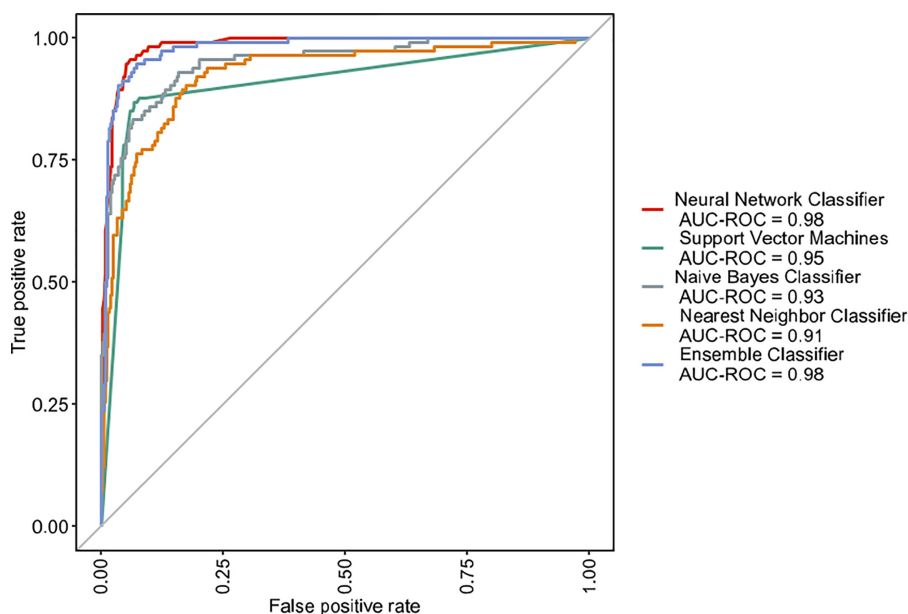


Figure 6.
ROC Curves of all models

models' capability to predict the true positive rate for all classes, and specificity measures the models' capability to predict the true negative rate for all classes.

The prediction results suggest that the accuracy of the ANNC is the highest with a value of 88%. According to the value of AUC, ANNCs and ECs have the highest value of 96%. Sensitivity and specificity results indicate that ANNCs have the highest predictive power. Different studies of [Ioannidis *et al.* \(2010\)](#), [Kim and Kang \(2010\)](#), [Yeh *et al.* \(2010\)](#), [Olson *et al.* \(2012\)](#), [Geng *et al.* \(2015\)](#) also found significant results by using ANNs. The result of kappa is lower for the NBC which indicates that NBC has moderate significance in the prediction of financial solvency. Nevertheless, the logLoss value EC outperformed all other models as its value is close to zero. Many researchers also found consistent results by applying EC results ([Heo, 2014](#); [Kim and Kang, 2010](#); [Pisula, 2020](#)). Also, [Figure 6](#) of ROC curves indicates the nearest neighbor classifier has a lower AUC. However, ANNCs revealed better outcomes for financial solvency prediction.

5. Implication and conclusions

Financial solvency prediction is linked with credit and liquidity risk, which were derived from the financial crisis and recent financial scams. Every investor wants to know where they are investing and what is their financial condition. However, there was a gap in the literature regarding machine learning for financial solvency prediction in the context of Bangladesh. The study aims to fill this gap by developing machine learning models in the context of Bangladesh for financial solvency prediction. For this reason, this study analyses 244 Dhaka stock exchanges; publicly listed companies have been selected, and data have been collected from the 2015 to 2019 period. And two subsets of data have also been developed for training and testing datasets. The uniqueness of this study comes from the large dataset in the context of Bangladesh and considering time series in data. For machine learning model training, it is necessary to classify observation in different categorized classes, so data are classified as

secure, healthy and insolvent by the analysis of the Altman Z-score. The next step of machine learning model building is feature selection. By reviewing the previous literature, a total of 11 features have been selected for model training. After selecting the model training and testing, all model performances are measured. This study found that the ANNC is the best model for financial solvency prediction among all model, EC also has consistent results (Heo, 2014; Kim and Kang, 2010; Pisula, 2020; Ioannidis *et al.*, 2010; Kim and Kang, 2010; Yeh *et al.*, 2010; Olson *et al.*, 2012).

The findings of this study can be used by academicians, investors and financial practitioners. The implication of this study is to develop a systematic way to implement the machine learning model in the enterprise resource planning software of companies. By doing so, decision-makers can assess their financial position. Bank, nonbank financial institutions and microcredit organizations can implement machine learning models for client categorization. Financial intermediaries can train predictive models with their client data and use that for loan sanctioning which will reduce the cost of loan sanctioning as well as credit card issuance. Finally, this study aims to make a guideline for the rational investor to decide before making any investment decision. This objective led the author to train the model with the full dataset of 1,200 observations and make the models available for future prediction in the context of Bangladesh. This predictive model is compiled through the *R* statistical programming language and named "Financial Solvency Prediction by Machine Learning Web App" and made available for the rational investor (Abdullah, 2020). By inputting all features, it will show the prediction results with their probabilities. This application can be used by all level users, from decision-makers to rational investors.

The major drawback is not considering external effects such as GDP, inflation rates, foreign exchange volatility, etc. along with the internal factors such as board composition, top management compensation, etc. Another drawback of this study is to only consider 11 features, while a larger number of features can improve the predictive power of the models. The outcome of this study should be implemented by every financial institution as machine learning can be used for credit scoring, credit rating, loan classification, etc. As prior experience is needed for building a machine learning predictive model, an individual-level investor can use the application developed by the author for investment decision-making including creating term deposits into the financial institution, purchasing a share or buying an insurance policy, etc. For experts, the results of this study are interesting as data transformations are not done but the prediction power was highly significant for the models. Future research can eliminate the limitations of this study and can consider external variables or add more features and tests to predict the credit rating, small and medium enterprise loan default classification, credit card default classification, corporate loan default prediction, etc.

References

- Abdullah, M. (2015), "An empirical analysis of liquidity, profitability and solvency of Bangladeshi banks", *Journal of Business and Financial Affairs*, Vol. 5, p. 157.
- Abdullah, M. (2020), "Financial solvency prediction by machine learning Web App", [Online], available at: <https://mabdullah.shinyapps.io/Financial-Solvency-Prediction/>.
- Altman, E.I. (1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, Vol. 23, pp. 589-609.
- Barboza, F., Kimura, H. and Altman, E. (2017), "Machine learning models and bankruptcy prediction", *Expert Systems with Applications*, Vol. 83, pp. 405-417.
- Beaver, W.H. (1966), "Financial ratios as predictors of failure", *Journal of Accounting Research*, Vol. 4, pp. 71-111.

- Booth, A., Gerding, E. and McGroarty, F. (2014), "Automated trading with performance weighted random forests and seasonality", *Expert Systems with Applications*, Vol. 41, pp. 3651-3661.
- Cecchini, M., Aytug, H., Koehler, G.J. and Pathak, P. (2010), "Making words work: using financial text as a predictor of financial events", *Decision Support Systems*, Vol. 50, pp. 164-175.
- Chandra, D.K., Ravi, V. and Bose, I. (2009), "Failure prediction of dotcom companies using hybrid intelligent techniques", *Expert Systems with Applications*, Vol. 36, pp. 4830-4837.
- Chen, M.-Y. (2011), "Predicting corporate financial distress based on integration of decision tree classification and logistic regression", *Expert Systems with Applications*, Vol. 38, pp. 11261-11272.
- Chen, H.L., Yang, B., Wang, G., Liu, J., Xu, X., Wang, S.J. and Liu, D.Y. (2011), "A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method", *Knowledge-Based Systems*, Vol. 24, pp. 1348-1359.
- Danenas, P. and Garsva, G. (2015), "Selection of support vector machines based classifiers for credit risk domain", *Expert Systems with Applications*, Vol. 42, pp. 3194-3204.
- du Jardin, P. and Séverin, E. (2012), "Forecasting financial failure using a Kohonen map: a comparative study to improve model stability over time", *European Journal of Operational Research*, Vol. 221, pp. 378-396.
- Geng, R., Bose, I. and Chen, X. (2015), "Prediction of financial distress: an empirical study of listed Chinese companies using data mining", *European Journal of Operational Research*, Vol. 241, pp. 236-247.
- Heo, J. and Yang, J.Y. (2014), "AdaBoost based bankruptcy forecasting of Korean construction companies", *Applied Soft Computing*, Vol. 24, pp. 494-499.
- Hosaka, T. (2019), "Bankruptcy prediction using imaged financial ratios and convolutional neural networks", *Expert Systems with Applications*, Vol. 117, pp. 287-299.
- Ioannidis, C., Pasiouras, F. and Zopounidis, C. (2010), "Assessing bank soundness with classification techniques", *Omega*, Vol. 38, pp. 345-357.
- Jabeur, S.B., Sadaoui, A., Sghaier, A. and Aloui, R. (2020), "Machine learning models and cost-sensitive decision trees for bond rating prediction", *Journal of the Operational Research Society*, Vol. 71, pp. 1161-1179.
- Jones, S. (2017), "Corporate bankruptcy prediction: a high dimensional analysis", *Review of Accounting Studies*, Vol. 22, pp. 1366-1422.
- Kim, M.-J. and Kang, D.-K. (2010), "Ensemble with neural networks for bankruptcy prediction", *Expert Systems with Applications*, Vol. 37, pp. 3373-3379.
- Kim, H.-J., Jo, N.-O. and Shin, K.-S. (2016), "Optimization of cluster-based evolutionary undersampling for the artificial neural networks in corporate bankruptcy prediction", *Expert Systems with Applications*, Vol. 59, pp. 226-234.
- Landis, J.R. and Koch, G.G. (1977), "An application of hierarchical Kappa-type statistics in the assessment of majority agreement among multiple observers", *Biometrics*, International Biometric Society, pp. 363-374.
- Liang, D., Lu, C.-C., Tsai, C.-F. and Shih, G.-A. (2016), "Financial ratios and corporate governance indicators in bankruptcy prediction: a comprehensive study", *European Journal of Operational Research*, Vol. 252, pp. 561-572.
- Mai, F., Tian, S., Lee, C. and Ma, L. (2019), "Deep learning models for bankruptcy prediction using textual disclosures", *European Journal of Operational Research*, Vol. 274, pp. 743-758.
- Masmoudi, K., Abid, L. and Masmoudi, A. (2019), "Credit risk modeling using Bayesian network with a latent variable", *Expert Systems with Applications*, Vol. 127, pp. 157-166.
- Michie, D., Spiegelhalter, D.J. and Taylor, C.C. (1994), "Machine learning", *Neural and Statistical Classification*, Vol. 13, pp. 1-298.

-
- Ohlson, J.A. (1980), "Financial ratios and the probabilistic prediction of bankruptcy", *Journal of Accounting Research*, pp. 109-131.
- Olson, D.L., Delen, D. and Meng, Y. (2012), "Comparative analysis of data mining methods for bankruptcy prediction", *Decision Support Systems*, Vol. 52, pp. 464-473.
- Pisula, T. (2020), "An ensemble classifier-based scoring model for predicting bankruptcy of polish companies in the Podkarpackie Voivodeship", *Journal of Risk and Financial Management*, Vol. 13 No. 2, p. 37.
- Raghupathi, W., Schkade, L.L. and Raju, B.S. (1991), "A neural network application for bankruptcy prediction", *Proceedings of the Twenty-Fourth Annual Hawaii International Conference on System Sciences*, Vol. 4, pp. 147-155.
- Shapiro, J. (2001), "Genetic algorithms in machine learning", in Paliouras, G., Karkaletsis, V. and Spyropoulos, C.D. (Eds), *Machine Learning and its Applications: Advanced Lectures*, Springer Berlin Heidelberg, Berlin (Heidelberg), pp. 146-168.
- Shi, Y. and Li, X. (2009), "A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms", *Heliyon*, Vol. 5 No. 12.
- Son, H., Hyun, C., Phan, D. and Hwang, H.J. (2019), "Data analytic approach for bankruptcy prediction", *Expert Systems with Applications*, Vol. 138, p. 112816.
- Sun, L. and Shenoy, P.P. (2007), "Using Bayesian networks for bankruptcy prediction: Some methodological issues", *European Journal of Operational Research*, Vol. 180, pp. 738-753.
- Tam, K. (1991), "Neural network models and the prediction of bank bankruptcy", *Omega*, Vol. 19 No. 5, pp. 429-445.
- Tavana, M., Abtahi, A.-R., Caprio, D.D. and Poortarigh, M. (2013), "An Artificial Neural Network and Bayesian Network model for liquidity risk assessment in banking", *Neurocomputing*, Vol. 275, pp. 2525-2554.
- Xiao, H., Xiao, Z. and Wang, Y. (2016), "Ensemble classification based on supervised clustering for credit scoring", *Applied Soft Computing*, Vol. 43, pp. 73-86.
- Yeh, C.-C., Chi, D.-J. and Hsu, M.-F. (2010), "A hybrid approach of DEA, rough set and support vector machines for business failure prediction", *Expert Systems with Applications*, Vol. 37, pp. 1535-1541.

Corresponding author

Mohammad Abdullah can be contacted at: htasfiq@gmail.com; si3861@putra.unisza.edu.my