Malaysian Journal of Computing, 7 (1): 1011-1026, 2022

eISSN: 2600-8238 online

TOWARDS DEVELOPMENT OF ROBUST MACHINE LEARNING MODEL FOR MALAYSIAN CORPORATION: A SYSTEMATIC REVIEW OF ESSENTIAL ASPECTS FOR CORPORATE CREDIT RISK ASSESSMENT

Zulkifli Halim[1] and Shuhaida Mohamed Shuhidan[2*]

[1]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Cawangan Pahang, Kampus Jengka*

[2]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam*

[1]zulkiflihalim@uitm.edu.my, [2*]shuhaida6704@uitm.edu.my

ABSTRACT

This study aims to review the essential aspects of credit risk assessment. The scope of this study is the credit risk assessment studies that used machine learning or used Malaysian data. This study is an overview of the development of robust machine learning for Malaysian corporation credit risk assessment. This study used a systematic review as the methodology. After thorough searching, this study has selected 20 studies to be reviewed. As a result, three essential aspects are identified: the variables, the features, and the methods used for financial distress prediction. This study found that financial ratios, macroeconomics, and corporate governance indicators are essential in credit risk assessment studies. The debt ratio was recorded as the most widely used ratio, found in 14 studies, followed by the liquidity ratio, used in 12 studies. In addition, the studies performed using Malaysian data show that the debt ratio and liquidity ratio are significant. Support vector machine (SVM) and genetic algorithm (GA) are among the best methods to be used. Recurrent neural network (RNN) is the latest credit risk assessment method to solve the time series data problem. In conclusion, all the essential aspects identified in this study should be considered in any credit risk assessment study.

Keywords: *Credit Risk Assessment, Machine Learning, Malaysian Corporation, Systematic Review.*

Received for review: 31-12-2021; Accepted: 23-02-2022; Published: 31-03-2022

1. Introduction

Credit risk assessment is the procedure by which investors or lenders predict the chances of loan default to measure risk (Azayite & Achchab, 2016; Tsai, 2014; Wang *et al.*, 2014). A wrong decision places the institution at risk (Babič *et al.*, 2013; Wang *et al.*, 2014). Credit risk assessment studies frequently employ the words: bankruptcy prediction, financial distress prediction, business failure prediction, and default prediction (Geng *et al.*, 2015). From a corporation's perspective, an organization is labelled either a healthy or non-healthy company based on the risk assessment (Zięba *et al.*, 2016). There is the possibility that a non-healthy organization leads to a failure and significantly influences economic growth (Chou *et al.*, 2017). The finance industry aims to invest in producing accurate analytical prediction tools (Antunes *et al.*, 2017). Analytical tools for credit risk assessment are in high demand because developing a better predictive model is very important. This has led researchers and practitioners to pay attention to these issues (Wang *et al.*, 2014).

Credit risk assessment depends on the financial indicators that represent the companies' status at a given time. The most common financial indicator used is the financial ratio. Based on the literature, lots of features have been used. The most used features are the financial ratio variables (Chou *et al.*, 2017; Du Jardin, 2017). Altman, a pioneer in credit risk assessment studies, used five financial ratios: profitability ratio, productivity ratio, liquidity ratio, leverage ratio, and activity ratio (Altman, 1968). Several studies have combined the financial ratio with other variables such as macroeconomics and corporate governance. For instance, Alifiah (2014) and Bakar *et al.* (2012) have used financial ratio and macroeconomic as the variables. Meanwhile, Abdullah *et al.* (2016), Liang *et al.* (2016), and Adiana *et al.* (2015) have used financial ratio and corporate governance as the variables.

Many methods for credit risk assessment are based on statistical (Achim *et al.*, 2012; Almamy *et al.*, 2016; Brîndescu-Olariu, 2017) and machine learning techniques (Azayite & Achchab, 2016; Chou *et al.*, 2017; Fallahpour *et al.*, 2017; Zelenkov *et al.*, 2017). Based on the past literature, most of the studies performed a comparison between statistical and machine learning techniques. Machine learning techniques can be a single classifier, an ensemble, or a hybrid method. As a result, many studies have shown that machine learning outperforms statistical techniques (Chou *et al.*, 2017; Du Jardin, 2017; Kwon *et al.*, 2017; Uthayakumar *et al.*, 2017; Zelenkov *et al.*, 2017).

The most common statistical techniques used are multivariate discriminant analysis (MDA) (Barboza *et al.*, 2017; Du Jardin, 2017; Kwon *et al.*, 2017) and logistic regression (LR) (Barboza *et al.*, 2017; Du Jardin, 2017; Kwon *et al.*, 2017; Uthayakumar *et al.*, 2017). Meanwhile, the most machine learning techniques used are neural network (NN) (Barboza *et al.*, 2017; Chou *et al.*, 2017; Du Jardin, 2017; Kwon *et al.*, 2017; Uthayakumar *et al.*, 2017; Zelenkov *et al.*, 2017), support vector machine (SVM) (Barboza *et al.*, 2017; Du Jardin, 2017; Fallahpour *et al.*, 2017; Kwon *et al.*, 2017; Uthayakumar *et al.*, 2017), decision tree (DT) (Du Jardin, 2017; Uthayakumar *et al.*, 2017; Zelenkov *et al.*, 2017), genetic algorithm (GA) (Chou *et al.*, 2017; Min, 2016; Zelenkov *et al.*, 2017), and k-nearest neighbour (kNN) (Liang *et al.*, 2016; Min, 2016), and others. In addition, the recent technique that was applied in credit risk assessment study is a recurrent neural network (RNN) (Kwon *et al.*, 2017), which is one of the deep learning techniques. Deep learning can solve many human problems easier even in a very complex situation (Exastax, 2018). RNN is a neural network application that can capture previous information that has been calculated to be used in the following sequence (Britz, 2015).

Corporation bankruptcy rates in Malaysia have shown an increase from 1998 to 2015. According to Trading Economics, by the end of the first quarter of this year, bankruptcy companies in Malaysia are expected to total 1874. By 2020, on average, the number of bankruptcy companies is projected to be 1850 companies per month (Trading Economics, 2018). This value has shown that Malaysia's credit risk assessment approach is inversely proportional to the number of corporate bankrupts. The factor may be due to weaknesses in the current approach. From the research perspective, most of the studies on credit risk assessment use Malaysian data using statistical techniques such as Abdullah *et al.* (2016), Adiana *et al.* (2015), Alifiah (2014), and several apply machine learning, such as Ramakrishnan *et al.* (2015).

The primary purpose of this paper is to identify the essential aspects of credit risk assessment for the corporation. This study will answer the following research questions:

- How many aspects are essential for corporate credit risk assessment?
- What are the components of each aspect for corporation credit risk assessment?

The scope of this study is the credit risk assessment studies that used machine learning or used Malaysian data. This study is an overview of the development of robust machine learning for Malaysian corporations' credit risk assessment using machine learning algorithms. The rest of this paper is organized as follows: Section 2 consists of the methodology used in this study.

The discussion on the three essential aspects of credit risk assessment studies is in Section 3. Then, Section 4 discussed the results of the review. Lastly, the conclusions are in Section 5.

2. Methodology

This study is used a systematic review methodology to review the related elements for corporation credit risk assessment, such as the variables, the features, and the methods used. Apart from that, the vital machine learning development tasks will also be reviewed. A systematic review is a method to produce valid and reliable knowledge with less bias and the required breadth of literature (Alaka *et al.*, 2018). According to Alaka *et al.* (2018), Google Scholar is one of the databases available to perform the literature search. Hence, this study uses Google Scholar databases to search for related literature. The defined string to perform the literature search in this study is ("forecasting" OR "prediction") AND ("bankruptcy" OR "insolvency" OR "distress" OR "default" OR "business failure") AND ("Malaysia" OR "machine learning"). The string defined is applied based on the (Alaka *et al.*, 2018) approach with the amendment. The keywords Malaysia and machine learning are used as the defined string since this study only wants to review machine learning methods or the studies that used Malaysian data. The years have been refined from 2014 to 2017 only.

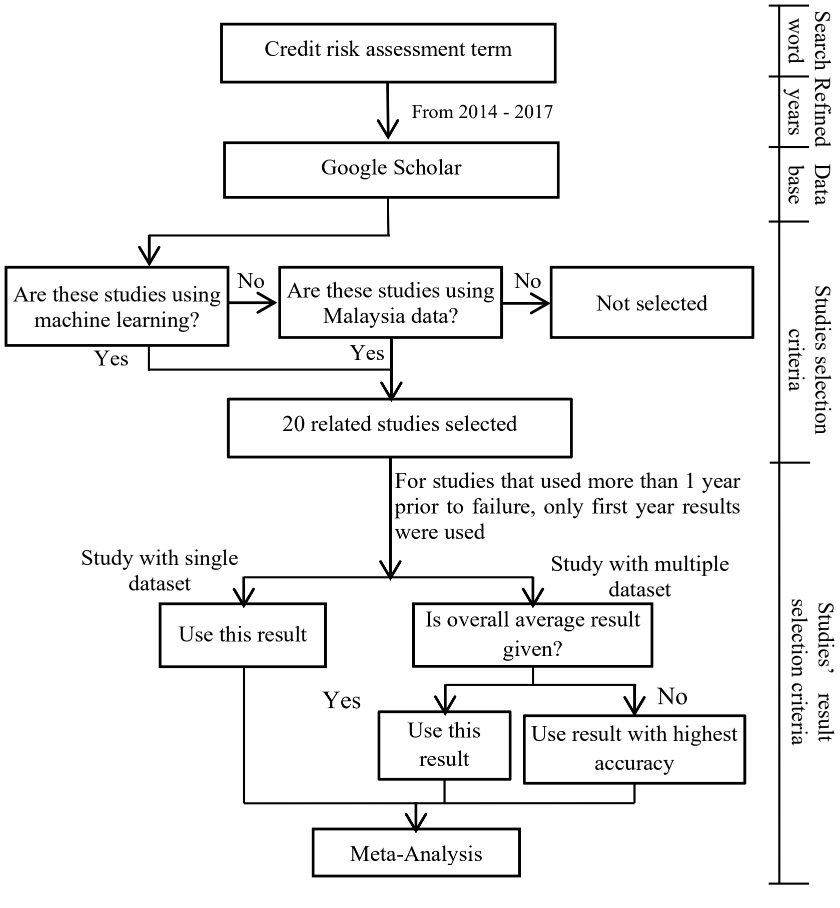

Figure 1. Process flow of the methodology (adapted based on (Alaka et al., 2018)).

Based on the search results, only 20 related research were selected that fulfil the requirements of this study without ambiguous reporting of results. Several studies are using more than one year of failure data. If so, only the results of the first-year of data are considered. A few studies also use multiple datasets to train and test their methods. To simplify it, use the results if the overall results are given. Otherwise, this study only uses the result with the highest accuracy. This methodology approach was adapted from Alaka *et al.* (2018). Finally, the meta-analysis is produced. The process flow of the methodology of this study is shown in Figure 1.

3. Essential aspects of credit risk assessment studies

This section describes the essential aspects of credit risk assessment studies. After a comprehensive review of the selected studies, three essential aspects have been identified. The three aspects identified are as follows:

A. Variables

There are three standard variables used for credit risk assessment studies. The variables are financial ratios (FR), macroeconomics (ME), and corporate governance (CG). FR is the ratio generated from a company's financial statements, such as balance sheets and income statements. FR has the financial condition insights (Yeh *et al.*, 2014). Macroeconomics is the variable that represents the economic situation of a country. According to Nouri & Soltani (2016), macroeconomics can improve the predictive power. Recently, many studies have shown that corporate governance indicators also played an essential role in credit risk assessment studies (Liang *et al.*, 2016). FR is the compulsory variable for credit risk assessment since it has been used in all studies. Several studies have combined with ME or CG or both.

B. Features

Various features have been used in different studies. The most commonly used features are categorized as FR variables. Generally, there are hundreds of FR features that can be generated. Usually, the features are grouped under a few ratios, such as profitability, productivity, solvency, liquidity, leverage, activity, and others (Fallahpour *et al.*, 2017). It is impossible to use all the FR features since many possible ratios could be produced. Only the ratios that are widely used are considered. The most used features under ME include gross domestic products (GDP) and price consumer index (CPI). The features of the CG indicator can be the age of the company (Abdullah *et al.*, 2016; Adiana *et al.*, 2015), the number of directors (Abdullah *et al.*, 2016; Liang *et al.*, 2016), and controlling shareholder (Abdullah *et al.*, 2016; Liang *et al.*, 2016).

C. Methods

There are lots of methods for credit risk assessment. Most methods are based on statistical (Achim *et al.*, 2012; Almamy *et al.*, 2016; Brîndescu-Olariu, 2017) and machine learning techniques (Azayite & Achchab, 2016; Chou *et al.*, 2017; Fallahpour *et al.*, 2017; Zelenkov *et al.*, 2017). Based on the literature, most of the studies performed a comparison of the credit risk assessment methods between statistical and machine learning approaches. As a result, many studies have shown that machine learning outperforms statistical techniques (Chou *et al.*, 2017; Du Jardin, 2017; Kwon *et al.*, 2017; Uthayakumar *et al.*, 2017; Zelenkov *et al.*, 2017).

The results of this study will be focused on all the essential aspects as discussed above. The results are discussed in the following sections.

4. Results and discussions

This section presents the results, analysis, and discussion of the systematic review summary represented as tables and charts. The results are presented concerning each identified criterion. Overall, the findings of this study showed that 16 studies had used machine learning (without using Malaysian data), three studies using Malaysian data (without using machine learning), and one used machine learning with Malaysian data.

Figure 2 shows the overall percentange of studies using machine learning and Malaysian data.

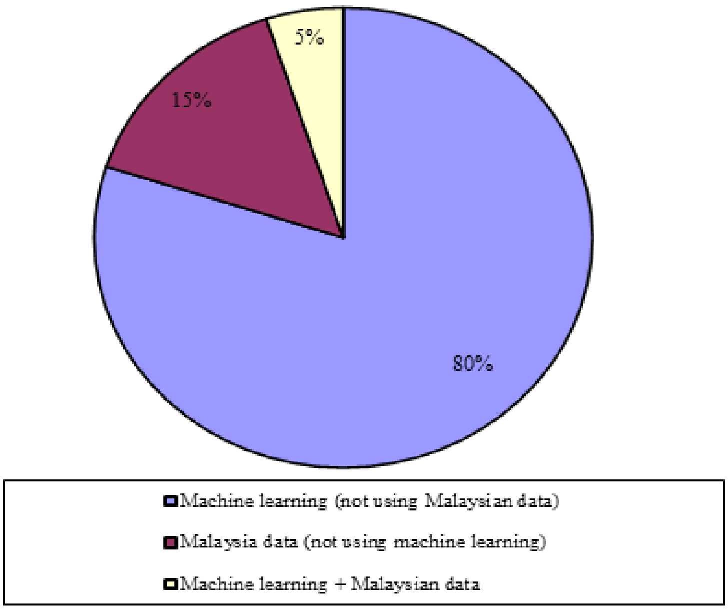

Figure 2. The overall percentage of the studies using machine learning and Malaysian data.

4.1. Variables

Three variables are commonly used for credit risk assessment studies based on the previous study. The variables are financial ratios (FR), macroeconomic (ME), and corporate governance (CG) indicators. Table 1 shows the variables and the datasets used in the previous studies. FR is an essential variable since it has been used in all studies. Several researchers have combined FR with other variables. For instance, Zelenkov *et al.* (2017) used all three variables and produced a better result with 93.4% accuracy. However, Zelenkov *et al.* (2017) did not investigate the effect of each variable further. Meanwhile, Alifiah (2014) used the FR with ME variables. Abdullah *et al.* (2016), Adiana *et al.* (2015), Chou *et al.* (2017), and Liang *et al.* (2016) have combined FR with CG. Figure 3 shows the percentages of the variables used for the credit risk assessment.

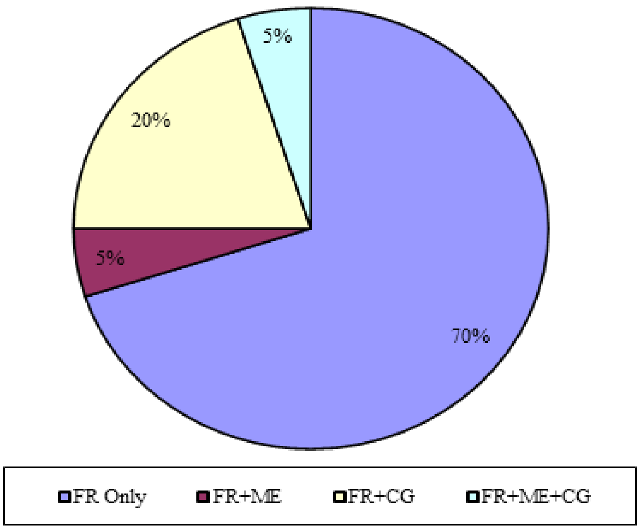

Figure 3. Percentage of the variables used for credit risk assessment.

Table 1. The variables and datasets used in the previous study.

References	Variables			Datasets		Results
	FR	ME	CG	Country	No of Features	(%)
(Zięba *et al.*, 2016)	✓	-	-	Polish	64	95.9*
(Chou *et al.*, 2017)	✓	-	✓	Taiwan	64	90.71
(Azayite & Achchab, 2016)	✓	-	-	Moroccan	5	84.5
(Wang *et al.*, 2014)	✓	-	-	Unknown	23	86.79
(Du Jardin, 2017)	✓	-	-	French	32	82.9*
(Alifiah, 2014)	✓	✓	-	Malaysia	5	85
(Abdullah *et al.*, 2016)	✓	-	✓	Malaysia	13	93.6
(Liang *et al.*, 2016)	✓	-	✓	Taiwan	190	81.5
(Adiana *et al.*, 2015)	✓	-	✓	Malaysia	8	84.1
(Fallahpour *et al.*, 2017)	✓	-	-	Iran	29	95.55
(Zelenkov *et al.*, 2017)	✓	✓	✓	Russia	55	93.4
(Kwon *et al.*, 2017)	✓	-	-	Korea	5	82.8
(Barboza *et al.*, 2017)	✓	-	-	America & Canada	11	87.06
(Min, 2016)	✓	-	-	Korea	75	76.18
(Ramakrishnan *et al.*, 2015)	✓	-	-	Malaysia	8	86.83
(Du Jardin, 2016)	✓	-	-	French	35	84.65
(Sun *et al.*, 2017)	✓	-	-	China	44	87.6
(Geng *et al.*, 2015)	✓	-	-	China	31	78.8
(Tsai *et al.*, 2014)	✓	-	-	Japan	15	88.36
(Heo & Yang, 2014)	✓	-	-	Korea	12	78.5

Most of the studies only used FR. Many studies showed that the combination of FR with other variables was able to produce more than 80% accuracy in results, such as (Abdullah *et al.*, 2016; Adiana *et al.*, 2015; Alifiah, 2014; Chou *et al.*, 2017; Liang *et al.*, 2016; Zelenkov *et al.*, 2017). Generally, the combination of the variables can produce a relatively good result for credit risk assessment. Most of the studies used accuracy percentage as the performance measurement. However, several studies used different measurements, which is the area under the curve (AUC) such as Du Jardin (2017), Zięba, *et al.*, (2016) as star labelled (*) in Table 1.

Four studies have used Malaysian corporation data (Abdullah *et al.*, 2016; Adiana *et al.*, 2015; Alifiah, 2014; Ramakrishnan *et al.*, 2015). In their studies, Abdullah *et al.* (2016), Adiana *et al.* (2015), and Alifiah (2014) have applied statistical techniques as the approach in their studies; meanwhile Ramakrishnan *et al.* (2015) have used machine learning. Figure 4 shows the number of studies that used different countries' data.

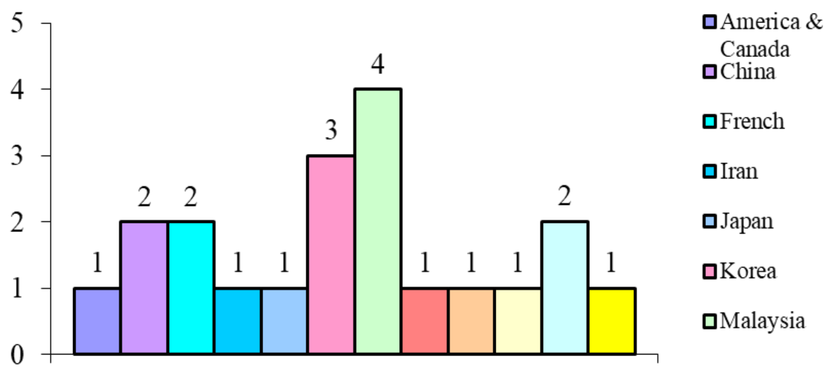

Figure 4. Number of studies that used different countries' data.

4.2. Features

Various features have been used in credit risk assessment studies. Table 2 shows the features that were primarily used in the previous study. The listed features for the FR variable are based on the number of features used in four or more studies. The most used features for FR are liquidity ratio, debt ratio, working capital / total asset, and return on total assets used in 8 or more studies. Debt ratio becomes the highest number of features used which has been used in 14 studies. However, for ME and CG indicators, the listed features are based on the number of features that were used in 2 or more studies since only a few studies used ME and CG indicators. There are two commonly used features for the ME variable: gross domestic product (GDP) and consumer price index (CPI). Meanwhile, the company's age, number of directors, and controlling shareholders are the features under the CG indicator used in the previous studies.

The lowest number of features used is five (Alifiah, 2014; Azayite & Achchab, 2016; Kwon *et al.*, 2017). Meanwhile, the highest number used was 190 (Liang *et al.*, 2016). Six studies have used more than 40 features, and five have used ten features or less. Figure 5 shows the range of the number of features used. Studies conducted using Malaysian data (Alifiah, 2014) have used five features, Adiana *et al.* (2015) and Ramakrishnan *et al.* (2015) have used eight features, and Abdullah *et al.* (2016) have used 13 features.

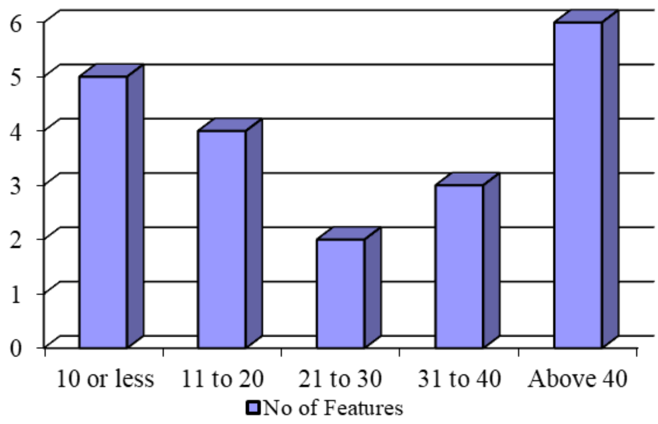

Figure 5. Number of features used for credit risk assessment.

Table 2. The list of features used in previous studies.

Variables	Features	Citations
Financial Ratio	Sales / Total Assets	(Abdullah *et al.*, 2016; Adiana *et al.*, 2015; Azayite & Achchab, 2016; Du Jardin, 2017; Heo & Yang, 2014; Wang *et al.*, 2014; Zięba *et al.*, 2016)
	Net Income / Total Assets	(Alifiah, 2014; Geng *et al.*, 2015; Kwon *et al.*, 2017; Liang *et al.*, 2016; Wang *et al.*, 2014; Zięba *et al.*, 2016)
	Working Capital / Total Asset	(Azayite & Achchab, 2016; Barboza *et al.*, 2017; Heo & Yang, 2014; Kwon *et al.*, 2017; Liang *et al.*, 2016; Wang *et al.*, 2014; Zelenkov *et al.*, 2017; Zięba *et al.*, 2016)
	Return on Asset	(Fallahpour *et al.*, 2017; Geng *et al.*, 2015; Liang *et al.*, 2016; Ramakrishnan *et al.*, 2015; Sun *et al.*, 2017; Zelenkov *et al.*, 2017)
	Return on Equity	(Adiana *et al.*, 2015; Chou *et al.*, 2017; Fallahpour *et al.*, 2017; Geng *et al.*, 2015)
	Profit margin or return on sales	(Fallahpour *et al.*, 2017; Geng *et al.*, 2015; Ramakrishnan *et al.*, 2015; Zelenkov *et al.*, 2017; Zięba *et al.*, 2016)
	Gross profit margin	(Fallahpour *et al.*, 2017; Liang *et al.*, 2016; Sun *et al.*, 2017; Zięba *et al.*, 2016)
	Return on Total Assets	(Adiana *et al.*, 2015; Azayite & Achchab, 2016; Barboza *et al.*, 2017; Chou *et al.*, 2017; Geng *et al.*, 2015; Heo & Yang, 2014; Ramakrishnan *et al.*, 2015; Zelenkov *et al.*, 2017; Zięba *et al.*, 2016)
	Retained earnings to total asset	(Azayite & Achchab, 2016; Barboza *et al.*, 2017; Liang *et al.*, 2016; Zelenkov *et al.*, 2017; Zięba *et al.*, 2016)
	Liquidity or Current Ratio or Working Capital Ratio	(Abdullah *et al.*, 2016; Adiana *et al.*, 2015; Alifiah, 2014; Chou *et al.*, 2017; Fallahpour *et al.*, 2017; Kwon *et al.*, 2017; Liang *et al.*, 2016; Min, 2016; Ramakrishnan *et al.*, 2015; Sun *et al.*, 2017; Zelenkov *et al.*, 2017; Zięba *et al.*, 2016)
	Acid Test or Quick Ratio	(Chou *et al.*, 2017; Fallahpour *et al.*, 2017; Liang *et al.*, 2016; Min, 2016; Zięba *et al.*, 2016)
	Debt to equity ratio	(Fallahpour *et al.*, 2017; Geng *et al.*, 2015; Liang *et al.*, 2016; Zelenkov *et al.*, 2017; Zięba *et al.*, 2016)
	Current assets to total assets	(Geng *et al.*, 2015; Heo & Yang, 2014; Liang *et al.*, 2016; Wang *et al.*, 2014; Zelenkov *et al.*, 2017)
	Cash flow from operations / Total debt	(Liang *et al.*, 2016; Min, 2016; Sun *et al.*, 2017; Wang *et al.*, 2014)
	Debt Ratio	(Abdullah *et al.*, 2016; Adiana *et al.*, 2015; Alifiah, 2014; Chou *et al.*, 2017; Du Jardin, 2016, 2017; Fallahpour *et al.*, 2017; Geng *et al.*, 2015; Kwon *et al.*, 2017; Liang *et al.*, 2016; Min, 2016; Ramakrishnan *et al.*, 2015; Zelenkov *et al.*, 2017; Zięba *et al.*, 2016)
	Equity Ratio	(Fallahpour *et al.*, 2017; Liang *et al.*, 2016; Min, 2016; Zięba *et al.*, 2016)
	Cash flow/Total sales	(Du Jardin, 2016, 2017; Liang *et al.*, 2016; Wang *et al.*, 2014)
	Total Assets Turnover ratio	(Alifiah, 2014; Chou *et al.*, 2017; Fallahpour *et al.*, 2017; Liang *et al.*, 2016; Ramakrishnan *et al.*, 2015; Sun *et al.*, 2017; Zelenkov *et al.*, 2017)
	Accounts Receivable Turnover	(Chou *et al.*, 2017; Fallahpour *et al.*, 2017; Geng *et al.*, 2015; Liang *et al.*, 2016; Zelenkov *et al.*, 2017)
	Net income/Total sales	(Du Jardin, 2016, 2017; Liang *et al.*, 2016; Min, 2016; Wang *et al.*, 2014)
	Net operating working capital/Total sales	(Du Jardin, 2016, 2017; Heo & Yang, 2014; Liang *et al.*, 2016; Wang *et al.*, 2014)
	Current assets/Total sales	(Du Jardin, 2016, 2017; Liang *et al.*, 2016; Wang *et al.*, 2014)
	Inventory Turnover	(Chou *et al.*, 2017; Fallahpour *et al.*, 2017; Geng *et al.*, 2015; Liang *et al.*, 2016; Wang *et al.*, 2014)
Macro-economic	Gross Domestic Product	(Alifiah, 2014; Zelenkov *et al.*, 2017)
	Consumer Price Index	(Alifiah, 2014; Zelenkov *et al.*, 2017)
Corporate Governance	Age of Company	(Abdullah *et al.*, 2016; Adiana *et al.*, 2015)
	Number of directors	(Abdullah *et al.*, 2016; Liang *et al.*, 2016)
	Controlling Shareholder	(Abdullah *et al.*, 2016; Liang *et al.*, 2016)

Table 3 shows the features used for credit risk assessment using Malaysian data. Liquidity and debt ratios have been used in all studies. It has shown that those features are significant for credit risk assessment studies.

Table 3. The most used features that based on Malaysian data.

Features	**(Alifiah, 2014)**	**(Abdullah *et al.*, 2016)**	**(Adiana *et al.*, 2015)**	**(Ramakrishnan *et al.*, 2015)**
Sales / Total Assets	-	✓	✓	-
Net Income / Total Assets	✓	-	-	-
Return on Asset	-	-	-	✓
Return on Equity	-	-	✓	-
Profit margin or return on sales	-	-	-	✓
Return on Total Assets	-	-	✓	✓
Liquidity Ratio	✓	✓	✓	✓
Debt Ratio	✓	✓	✓	✓
Total Assets Turnover ratio	✓	-	-	✓
Gross Domestic Product	✓	-	-	-
Consumer Price Index	✓	-	-	-
Age of Company	-	✓	✓	-
Number of directors	-	✓	-	-
Controlling Shareholder	-	✓	-	-

4.3. Methods

There are lots of methods used in the previous studies for credit risk assessment. Table 4 shows the methods that have been used in the previous studies. The methods have been divided into two categories: statistical and machine learning. Machine learning (ML) methods categorized into three: ML algorithms, ensembles, and hybrids. Figure 6 shows the percentage of studies that performed the comparisons between the methods in statistical and machine learning. Most of the studies performed comparisons between statistical methods and machine learning. As a result, many studies have shown that machine learning outperforms statistical techniques (Chou *et al.*, 2017; Du Jardin, 2017; Kwon *et al.*, 2017; Uthayakumar *et al.*, 2017; Zelenkov *et al.*, 2017). Several studies only performed comparisons only between machine learning methods. Meanwhile, a few studies performed comparisons between statistical data only.

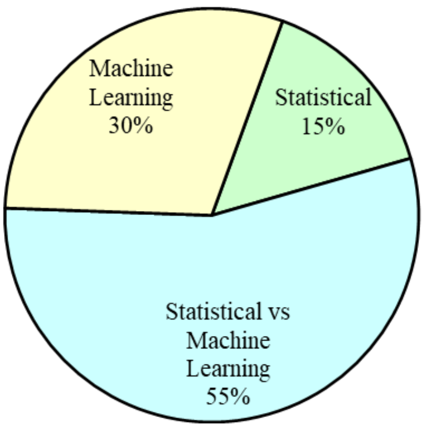

Figure 6. Percentage of studies performed comparisons between the methods among statistical and machine learning.

Table 4. The methods used for credit risk assessment studies.

Citations	Statistical	Machine Learning (ML)			The Best Method	Category	Results (%)
		ML Algorithm	Ensemble	Hybrid			
(Abdullah *et al.*, 2016)	LR	-	-	-	LR	LR	93.6
(Adiana *et al.*, 2015)	LR	-	-	-	LR	LR	84.1
(Alifiah, 2014)	LR	-	-	-	LR	LR	85
(Azayite & Achchab, 2016)	MDA	NN	-	NN + SOM	NN+SOM	NN (hybrid)	84.5
(Barboza *et al.*, 2017)	MDA, LR	NN, SVM, DT	BAG, BST	-	DT	DT	87.06
(Chou *et al.*, 2017)	-	NN, GA	-	GA + Fuzzy Clustering	GA+Fuzzy	GA (hybrid)	90.71
(Du Jardin, 2016)	MDA, LR	NN, DT	BST, BAG, RS	Classifier + Ensemble + PBM	PBM + RS +NN	NN (hybrid)	84.65
(Du Jardin, 2017)	MDA, LR, Cox's model	NN, SVM, DT, ELM	BAG, BST, RS, ROF, DEC	SOM + All statistical and classifiers SOM + all ensembles	BST+SOM+ ELM	Ensemble (hybrid)	82.9*
(Fallahpour *et al.*, 2017)	-	SVM	-	-	SVM with FS (SFFS)	SVM	95.55
(Geng *et al.*, 2015)	-	NN, SVM, DT, RST	MV	-	NN	NN	78.8
(Heo & Yang, 2014)	Z-score	MLP, SVM, DT	BST, BAG	-	BST (AB)	DT (ensemble)	78.5
(Kwon *et al.*, 2017)	MDA, LR	NN, SVM, RNN	-	-	RNN	RNN	82.8
(Liang *et al.*, 2016)	-	NN, SVM, DT, NB, kNN	-	-	SVM with FS (Stepwise DA)	SVM	81.5
(Min, 2016)	LR	SVM, kNN, GA	RS	GA-Based + Heterogeneo-us RS	GAHRS	GA (hybrid)	76.18
(Ramakrishnan *et al.*, 2015)	LR	NN, DT, SVM	BST, BAG	-	SVM + Bagging SVM + Boosting	SVM (ensemble)	86.83
(Sun *et al.*, 2017)	-	SVM	BST	AB-SVM +TimeBoost SVM, AB-SVM +Time weighting)	AB-SVM +Time weighting	SVM (hybrid)	87.60
(Tsai *et al.*, 2014)	-	NN, SVM, DT	BST, BAG	-	DT+BST	DT (ensemble)	88.36
(Wang *et al.*, 2014)	LR	SVM, DT, NB	BST, BAG	-	BST with FS	Ensemble	86.79
(Zelenkov *et al.*, 2017)	Z-score, LDA, QDA, LR	NN, DT, kNN, SVM, NB	BST, BAG, MV, XGB	Hybrid GA (TSCM)	Hybrid GA	GA (hybrid)	93.4
(Zięba *et al.*, 2016)	LDA, LR, JRip	NN, SVM, DT	BST, XGB	XGBE, EXGB	EXGB	DT (hybrid)	95.9*

AB=AdaBoost, DT=Decision Tree, BAG=Bagging, BST=Boosting, CS=Cost Sensitive, DEC=Decorate, ELM=Extreme Learning Machine, EXGB=Extreme boosted tree, FS=Feature Selection, GA=Genetic Algorithm, LDA=Linear DA, LR=Logistic Regression, kNN=k Nearest Neighbors, MDA=Multivariate Discriminant Analysis, MV=Majority Voting, NB=Naïve Bayes, NN=Neural Network, PBM=Profile Based Model, QDA=Quadratic Discriminant Analysis, RF=Random Forest, RNN=Recurrent NN, ROF=Rotation Forest, RST=Rough Set Theory, RS=Random Subspace, SFFS=Sequential Floating Forward Selection, SOM=Self-Organizing Maps, SVM=Support Vector Machine, TSCM=Two Step Classification Method, XGB=Extreme Gradient Boosting, XGBE=Ensemble of boosted trees

Multivariate discriminant analysis (MDA) and logistic regression (LR) are the most commonly used statistical techniques. While neural networks (NN), support vector machines (SVM), decision trees (DT), and genetic algorithms (GA) are the most commonly used machine learning algorithms. Besides that, the following techniques have also been used, such as k-nearest neighbour (kNN), rough set theory (RST), and naïve Bayes (NB). In addition, the recent technique that has been used in credit risk assessment studies is a recurrent neural network (RNN) (Kwon *et al.*, 2017), which is a deep learning technique. The ensemble is a technique in machine learning by combining ensemble methods such as boosting (BST), bagging (BAG), decorating (DEC), random subspace (RS), rotation forest (ROF), majority voting (MV), and extreme gradient boosting (XGB) with single classifiers. Meanwhile, the hybrid is a technique that combines multiple methods based on statistical, ML algorithm, and ensemble learning.

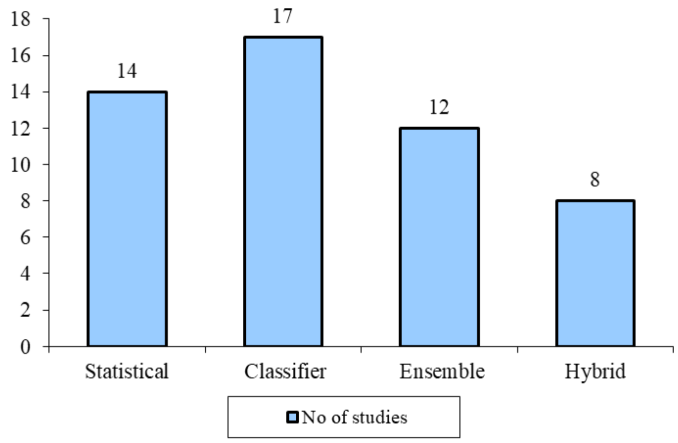

Figure 7. Number of studies that applied statistical and machine learning techniques.

Figure 7 shows the number of applied statistical and machine learning techniques in credit risk assessment studies. Table 4 shows the best methods based on the highest accuracy or AUC percentage in credit risk assessment: LR, DT, GA, NN, SVM, RNN, and ensemble (based on the category of the methods). Most of the studies used accuracy percentage as the performance measurement. However, several studies used the area under the curve (AUC), such as Du Jardin (2017) and Zięba *et al.* (2016), as star labelled (*) in Table 4. Three studies have shown that LR is the best method. However, those studies did not compare with other machine learning approaches. Figure 8 shows the number of studies for the best methods. In four studies, DT and SVM outperformed the other methods.

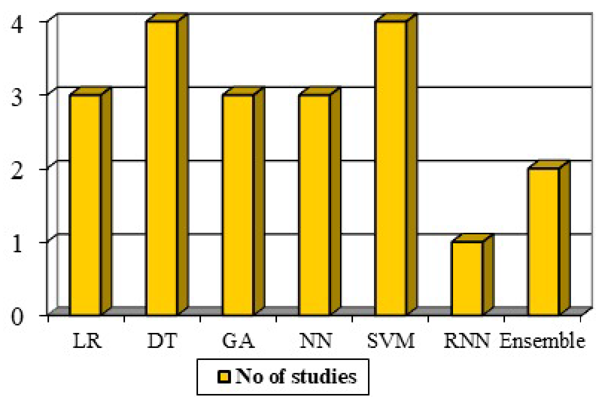

Figure 8. The number of studies for the best methods.

Overall, the methods that achieved the best results of more than 90% are LR with 93.6 % (Abdullah *et al.*, 2016), DT with 95.9 % (Zięba *et al.*, 2016), SVM with 95.55 % (Fallahpour *et al.*, 2017), and GA with 90.71 % (Chou *et al.*, 2017) and 93.4 % (Zelenkov *et al.*, 2017). The factors that influenced the results of studies performed by Abdullah *et al.* (2016), Chou *et al.* (2017), and Zelenkov *et al.* (2017) may be caused by the combination of FR with other variables. For instance, Zelenkov *et al.* (2017) have combined FR with ME and CG; meanwhile, Chou *et al.* (2017) and Abdullah *et al.* (2016) have combined FR with CG. On the other hand, the results obtained by Zięba *et al.* (2016) and Fallahpour *et al.* (2017) are probably due to the implementation of methods (hybrid) or the data (the features used) itself. Table 5 shows the average results of each best method. The highest average is put in boldface, while the lowest average is put in italic. From Table 4, the highest average of the results belongs to SVM with 87.87%. Meanwhile, the lowest average is NN with 82.65%.

Table 5. The average results of each best method.

Methods	Result 1 (%)	Result 2 (%)	Result 3 (%)	Result 4 (%)	Average (%)
LR	85	93.6	84.1	-	87.57
DT	95.9	87.06	88.36	78.5	87.45
GA	90.71	93.4	76.18	-	86.76
NN	84.5	84.65	78.8	-	*82.65*
SVM	81.5	95.55	86.83	87.6	**87.87**
RNN	82.8	-	-	-	82.8
Ensemble	86.79	82.9	-	-	84.85

In brief, this study has shown that SVM is the best method because of three reasons. The first reason is that the SVM algorithm can produce up to 95.55 %, performed by Fallahpour *et al.* (2017). Secondly, is because SVM achieved the highest average of results, with 87.87%. Thirdly, SVM has shown the best results in four studies (Fallahpour *et al.*, 2017; Liang *et al.*, 2016; Ramakrishnan *et al.*, 2015; Sun *et al.*, 2017). In addition, GA is another method that can produce a better result. Although GA is only outperformed by the other methods in three studies (Chou *et al.*, 2017; Min, 2016; Zelenkov *et al.*, 2017), but GA method can produce more than 90% results in two studies as performed by Chou *et al.* (2017) and Zelenkov *et al.* (2017).

4.4. Limitations of the current methods

Nowadays, time-series or sequence data poses new challenges for machine learning research. Many issues related to time series data need attention from the researchers, such as noise interference. Those issues can decrease the performance of machine learning (Awad & Khanna, 2016). In general, traditional GA and SVM are difficult to use with time-series data. Deep learning is a better approach to learning time series data (Wang et al, 2014; Yeh *et al.*, 2014). Financial data that has been used for credit risk assessment studies is an example of time-series data since the data was captured from the previous years prior to failure. Several studies only used a single year's data to predict failures, such as Heo and Yang (2014) and Zelenkov *et al.* (2017). Several studies have been used for several years prior to failure. However, most of the studies have evaluated and predicted each year separately (Azayite & Achchab, 2016; Chou *et al.*, 2017; Du Jardin, 2017; Geng *et al.*, 2015).

RNN has good performance for learning time series data over other techniques (Kwon *et al.*, 2017; Pascanu *et al.*, 2012). However, RNN is hard to train because of the vanishing gradient and exploding gradient problems (Du *et al.*, 2015; Jozefowicz *et al.*, 2015; Kwon *et al.*, 2017; Le *et al.*, 2015). There are many solutions to this problem, such as gating units, gradient clipping, better optimizers, and steeper gates. The most commonly used are gating units. There are two gating unit techniques: Long-Term Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). RNN with LSTM has been used to solve the credit risk prediction problem. According to Jozefowicz *et al.* (2015), GRU performed better than LSTM. The GRU was introduced by Cho *et al.* (2014). However, based on the literature in this study, the implementation of RNN with GRU has not been found in credit risk assessment studies.

5. Conclusion

In conclusion, the purpose of this study is to review the essential aspects of credit risk assessment studies. To answer the first research question, this study found that there are three main aspects of credit risk assessment. The aspects are the variables, features, and methods used for credit risk assessment studies. Next is to answer the second research question. The following paragraphs explain the components of each aspect of the corporation's credit risk assessment.

Three variables were primarily used in the previous study: financial ratio, macroeconomic, and corporate governance indicators. The financial ratio variable is used in all credit risk assessment studies. Several studies have also combined financial ratios with others, such as macroeconomic and corporate governance indicators. Combinations of those can improve the performance of credit risk assessment methods. Therefore, the macroeconomic and corporate governance variables should be considered in credit risk assessment studies with the financial ratio variable.

The second aspect is the features used for credit risk assessment. The liquidity ratio, debt ratio, working capital to total assets ratio, and return on total assets ratio are the metrics widely used in the previous studies. Those studies have been used in eight or more studies. The debt ratio was recorded as the most widely used ratio, found in 14 studies, followed by the liquidity ratio, which has been used in 12 studies. In addition, the studies using Malaysian data show that the debt ratio and liquidity ratio are significant for credit risk assessment in Malaysia. Gross domestic product (GDP) and price consumer index (CPI) are the most used features for macroeconomics variables. Meanwhile, three features are widely used for corporate governance: company age, number of directors, and controlling shareholder. Therefore, those stated features should be used in credit risk assessment studies.

The third aspect is the methods used for credit risk assessment. Many methods have been applied to credit risk assessment studies. The methods are either statistical or machine learning approaches. This study found that SVM and GA are among the best methods applied for credit risk assessment studies. Based on the review, the SVM method can produce an accuracy percentage of up to 95.5 % (Fallahpour *et al.*, 2017). Besides that, the average result shows that SVM achieved the highest percentage, which is 87.87%.

Another point that shows SVM is among the best methods is that SVM performed better than other methods in four studies (Fallahpour *et al.*, 2017; Liang *et al.*, 2016; Ramakrishnan *et al.*, 2015; Sun *et al.*, 2017). Another method, GA, is also considered among the best methods since GA can produce more than 90% accuracy in two studies (Chou *et al.*, 2017; Zelenkov *et al.*, 2017). Time-series data has become the new challenge for machine learning studies, especially in finance. RNN, the latest credit risk assessment method, has good performance for learning time series data compared to other techniques.

This study has three limitations. Firstly, this study did not review the essential tasks for the machine learning method. A lot of efforts need to be made for any machine learning method. Tasks such as feature selection, feature engineering, and parameter tuning are essential to produce an optimal machine learning model for credit risk assessment. Secondly, there is no discussion about the size of the data. Size data will also influence the performance of the methods for credit risk assessment. Lastly, time-series data should be discussed in detail. Therefore, as further work, all the limitations mentioned above will be considered.

Acknowledgment

The authors gratefully acknowledge Universiti Teknologi MARA and the Faculty of Computer and Mathematical Sciences for their support and resources.

References

Abdullah, N. A. H., Ma'aji, M. M., & Khaw, K. L.-H. (2016). The Value of Governance Variables in predicting Financial Distress among Small and Medium-Sized Enterprises in Malaysia. *Asian Academy of Management Journal of Accounting and Finance*, *12*(Suppl. 1), 75–88. https://doi.org/10.21315/aamjaf2016.12.S1.4.

Achim, M. V., Mare, C., & Borlea, S. N. (2012). A Statistical Model of Financial Risk Bankruptcy Applied for Romanian Manufacturing Industry. *Procedia Economics and Finance*, *3*(12), 132–137. https://doi.org/10.1016/S2212-5671(12)00131-1.

Adiana, N., Ahmad, A. H., Md. Rus, R., & Zainudin, N. (2015). Modelling Small Business Failures in Malaysia. *Social Science Research Network*, *February*, 1–10. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2402129.

Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S. O., Akinade, O. O., & Bilal, M. (2018). Systematic Review of Bankruptcy Prediction Models: Towards A Framework for Tool Selection. *Expert Systems with Applications*, 94, 164–184. https://doi.org/https://doi.org/10.1016/j.eswa.2017.10.040.

Alifiah, M. N. (2014). Prediction of Financial Distress Companies in the Trading and Services Sector in Malaysia Using Macroeconomic Variables. *Procedia - Social and Behavioral Sciences*, 129, 90–98. https://doi.org/10.1016/j.sbspro.2014.03.652.

Almamy, J., Aston, J., & Ngwa, L. N. (2016). An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Journal of Corporate Finance*, 36, 278–285. https://doi.org/10.1016/j.jcorpfin.2015.12.009.

Altman, E. I. (1968). The Prediction of Corporate Bankruptcy: A Discriminant Analysis. *The Journal of Finance*, 23(1), 193–194. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x/pdf.

Antunes, F., Ribeiro, B., & Pereira, F. (2017). Probabilistic modeling and visualization for bankruptcy prediction. *Applied Soft Computing Journal*, *In Press*. https://doi.org/10.1016/j.asoc.2017.06.043.

Azayite, F. Z., & Achchab, S. (2016). Hybrid Discriminant Neural Networks for Bankruptcy Prediction and Risk Scoring. *Procedia Computer Science*, 83(Ant), 670–674. https://doi.org/10.1016/j.procs.2016.04.149.

Babič, F., Havrilová, C., & Paralič, J. (2013). Knowledge discovery methods for bankruptcy prediction. *Lecture Notes in Business Information Processing*, 157, 151–162. https://doi.org/10.1007/978-3-642-38366-3_13.

Bakar, S., Kiong, I. T. W., & Nassir, A. M. (2012). Examining and validating a bankruptcy prediction models in Malaysia. *ICSSBE 2012 - Proceedings, 2012 International Conference on Statistics in Science, Business and Engineering: "Empowering Decision Making with Statistical Sciences,"* 219–223. https://doi.org/10.1109/ICSSBE.2012.6396560.

Barboza, F., Kimura, H., & Altman, E. I. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83, 405–417. https://doi.org/10.1016/j.eswa.2017.04.006.

Brîndescu-Olariu, D. (2017). Bankruptcy prediction logit model developed on Romanian paired sample Yearly bankruptcy frequency in Romania. *Theoretical and Applied Economics*, 24(1), 5–22.

Britz, D. (2015). *Recurrent Neural Networks Tutorial, Part 1 – Introduction to RNNs*. WILDML Artificial Intelligence, Deep Learning, and NLP.

Cho, K., van Merrienboer, B., Bahdanau, D., & Bengio, Y. (2014). *On the Properties of Neural Machine Translation: Encoder-Decoder Approaches*. https://doi.org/10.3115/v1/W14-4012

Chou, C. H., Hsieh, S. C., & Qiu, C. J. (2017). Hybrid genetic algorithm and fuzzy clustering for bankruptcy prediction. *Applied Soft Computing*, 56, 298–316. https://doi.org/10.1016/j.asoc.2017.03.014.

Du Jardin, P. (2016). A two-stage classification technique for bankruptcy prediction. *European Journal of Operational Research*, 254(1), 236–252. https://doi.org/10.1016/j.ejor.2016.03.008.

Du Jardin, P. (2017). Dynamics of firm financial evolution and bankruptcy prediction. *Expert Systems with Applications*, 75, 25–43. https://doi.org/10.1016/j.eswa.2017.01.016.

Exastax. (2018). *Comparison of Deep learning Frameworks*. Exastax Deep Learning Blog. https://www.exastax.com/deep-learning/a-comparison-of-deep-learning-frameworks/.

Fallahpour, S., Lakvan, E. N., & Zadeh, M. H. (2017). Using an ensemble classifier based on sequential floating forward selection for financial distress prediction problem. *Journal of Retailing and Consumer Services*, 34(March 2016), 159–167. https://doi.org/10.1016/j.jretconser.2016.10.002.

Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research*, 241(1), 236–247. https://doi.org/10.1016/j.ejor.2014.08.016.

Heo, J., & Yang, J. Y. (2014). AdaBoost based bankruptcy forecasting of Korean construction companies. *Applied Soft Computing Journal*, 24, 494–499. https://doi.org/10.1016/j.asoc.2014.08.009.

Kwon, H., Lee, D., & Shin, M. (2017). Dynamic forecasts of bankruptcy with Recurrent Neural Network model. *Journal of Intelligent Information Systems*, 23(3), 139–153.

Liang, D., Lu, C. C., Tsai, C. F., & Shih, G. A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research*, 252(2), 561–572. https://doi.org/10.1016/j.ejor.2016.01.012

Min, S. H. (2016). A genetic algorithm-based heterogeneous random subspace ensemble model for bankruptcy prediction. *International Journal of Applied Engineering Research*, 11(4), 2927–2931.

Nouri, B. A., & Soltani, M. (2016). Designing a bankruptcy prediction model based on account , market and macroeconomic variables. *Iranian Journal of Management Studies*, 9(1), 125–147.

Ramakrishnan, S., Mirzaei, M., & Bekri, M. (2015). Corporate default prediction with adaboost and bagging classifiers. *Jurnal Teknologi*, 73(2), 45–50.

Sun, J., Fujita, H., Chen, P., & Li, H. (2017). Dynamic financial distress prediction with concept drift based on time weighting combined with Adaboost support vector machine ensemble. *Knowledge-Based Systems*, *120*, 4–14. https://doi.org/10.1016/j.knosys.2016.12.019.

Trading Economics. (2018). *Malaysia Bankruptcies Forecast*. Trading Economics. https://tradingeconomics.com/malaysia/bankruptcies/forecast.

Tsai, C. F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion*, 16(1), 46–58. https://doi.org/10.1016/j.inffus.2011.12.001.

Tsai, C. F., Hsu, Y. F., & Yen, D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing Journal*, 24, 977–984. https://doi.org/10.1016/j.asoc.2014.08.047.

Uthayakumar, J., Vengattaraman, T., & Dhavachelvan, P. (2017). Swarm intelligence based classification rule induction (CRI) framework for qualitative and quantitative approach: An application of bankruptcy prediction and credit risk analysis. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2017.10.007.

Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, 41(5), 2353–2361. https://doi.org/10.1016/j.eswa.2013.09.033.

Yeh, C. C., Chi, D. J., & Lin, Y. R. (2014). Going-concern prediction using hybrid random forests and rough set approach. *Information Sciences*, 254, 98–110. https://doi.org/10.1016/j.ins.2013.07.011.

Zelenkov, Y., Fedorova, E., & Chekrizov, D. (2017). Two-step classification method based on genetic algorithm for bankruptcy forecasting. *Expert Systems with Applications*, 88, 393–401. https://doi.org/10.1016/j.eswa.2017.07.025.

Zięba, M., Tomczak, S. K., & Tomczak, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems with Applications*, *58*, 93–101. https://doi.org/10.1016/j.eswa.2016.04.001.

.