Capital Markets Review Vol. 28, No. 2, pp. 29-41 (2020)

Predicting SMEs Failure: Logistic Regression vs Artificial Neural Network Models

Juraini Zainol Abidin[1], Nur Adiana Hiau Abdullah[1]
& Karren Lee-Hwei Khaw[2,*]
[1]*College of Business, Universiti Utara Malaysia, Malaysia.*
[2]*Faculty of Business and Accountancy, University of Malaya, Malaysia.*

Abstract: Research questions: This study compares the power of logit and artificial neural network (ANN) models in predicting the failure of SMEs in the hospitality industry and identifies the predictors that are significant in determining business failure. **Motivation:** SMEs are an important segment of the Malaysian economy and contribute significantly to the country's economic growth. However, SMEs are riskier and associated with a high failure rate. In Malaysia, around 3.5% of the SMEs in the hospitality industry fail within the first two years and 54% of them cease operations within four. **Idea:** The use of ANN to model business failure, particularly in the hospitality industry, is relatively unexplored in the emerging markets. Based on the literature, this study hypothesizes that ANN models outperform logit models because of less stringent model assumptions. **Data:** Excluding missing information, a matched sample of 41 failed and 41 non-failed SMEs in the hospitality industry was identified from the year 2000 to 2016. The accounting ratios, firm-specific characteristics and governance variables are selected as potential predictors of SMEs failure in the hospitality industry. **Method/Tools:** Stepwise logit regression and multilayer perceptron ANN models were used to determine significant predictors to predict business failure. Each model's predictive power was compared. **Findings:** The ANN model was found to consistently outperform the logit model in classifying the failed and non-failed SMEs in the hospitality industry. Furthermore, the ANN model ranked liquidity as the most important predictor, followed by profitability and leverage, in determining business failure. Board size was also found to be a significant predictor in addition to the financial variables. The stepwise logit model also suggests a significant relationship between board size and the failure of SMEs. Therefore, in addition to financial predictors, a firm's governance is also key to business survival. **Contributions:** The findings of this study contribute to the limited literature on SMEs in the hospitality industry by providing empirical evidence from an emerging market perspective. The failure prediction model can be utilized to warn of potential business failure so that strategic measures can be taken to mitigate the risk of failure.

Keywords: Artificial neural network, business failure, hospitality industry, logistic regression, SMEs.
JEL classification: G30, G33

* Corresponding author: Karren Lee-Hwei Khaw. Tel.: 603-79673905. Email: karren@um.edu.my

Received 29 Mar 2020; Final revised 4 Nov 2020; Accepted 22 Sept 2020; Available online 8 Nov 2020.
To link to this article: https://www.mfa.com.my/cmr/v28_i2_a3/

1. Introduction

Studies on bankruptcy prediction models tend to emphasize public-listed firms. Studies on small and medium-sized enterprises (SMEs) are relatively limited due to the challenges in accessing the financial data and information of SMEs (Altman and Sabato, 2007). SMEs are important to the economy of every country. Malaysia is no exception, as SMEs make up 98.5% of the total business establishments and 65% of the total workforce (Department of Statistics Malaysia, 2017). SMEs recorded a gross domestic product (GDP) growth of 6.2% in 2018 (SME Corporation, 2019). The tourism industry is among the most important industries in Malaysia, which generates significant foreign exchange earnings and contributes significantly to the country's economic growth (Department of Statistics Malaysia, 2019). This growth affects the hospitality industry, which is categorized under the ambit of the tourism industry.

However, businesses within this industry are exposed to a relatively high failure rate. For example, in the United Kingdom, on average, 13.4% of the businesses in the accommodation and restaurant sector have failed each year between the observation period of 2012 to 2017 (Office for National Statistics, 2018). In Malaysia, around 3.5% of the SMEs in the hospitality industry fail within the first two years and 54% cease operations within four years (Shah and Ali, 2011). As a result, it is essential to have a model that can more closely predict business failure to provide an early warning of potential distress so that strategic measures can be taken well ahead to mitigate potential business failure.

This study aims to compare the prediction accuracy of logistic and multilayer perceptron artificial neural network (ANN) models in determining the failure of SMEs in the hospitality industry of Malaysia. The logistic regression or the logit model is the commonly used model to predict business failure, while the ANN model is a non-traditional model (Alaka *et al.*, 2018; Jackson and Wood, 2013). Studies have found that the ANN model can generate comparable or better classification results than the logit model (Kim, 2011). The key advantage of the ANN model over the other traditional models is its ability to recognize patterns regardless of the functional form (Charitou *et al.*, 2004; Khermkhan and Chancharat, 2013; Kim, 2011). An ANN can examine non-linear relationships and adjust the model adaptively (Sung *et al*., 1999).

However, the use of ANNs to model business failure in the hospitality industry is relatively unexplored in emerging markets, such as Malaysia. Therefore, to achieve the research objective, this study leverages the logit and ANN models to first determine the predictors that significantly classify failed and non-failed SMEs and then estimate the accuracy rate of each model to determine whether ANNs outperform logit models as hypothesized.

This study contributes to the literature in several ways. First of all, the findings contribute to the literature on bankruptcy prediction models for SMEs. This study shows that the multilayer perceptron ANN model does a better job, compared to the logit model, in classifying failed and non-failed SMEs. Second, this study provides evidence from the emerging market's perspective, such as Malaysia. Emerging markets are commonly perceived to be riskier than developed markets, as SMEs have to face stiff cross-border trade challenges and lack of access to finance and expertise that potentially increases the SMEs' probability of failure. Third, this study complements the studies by Abdullah *et al.* (2016a, 2016b) that model the failure of manufacturing SMEs in Malaysia.

The rest of the paper is organized as follows: Section 2 provides a review of the literature, Section 3 discusses the methodology, Section 4 presents the empirical results, and the final section concludes the paper.

2. Literature Review

Among the earliest works on bankruptcy prediction models are the studies by Beaver (1966), who used univariate analysis, and Altman (1968), who employed multiple discriminant analysis (MDA) models. Following Altman's work, several other business failure prediction models were introduced. These included the use of logit analysis by Ohlson (1980) and ANN by Odom and Sharda (1990).

The key benefit of the logit model is that it can handle non-linear relationships (Altman and Sabato, 2007). It is a conditional probability model that consists of a combination of variables that best discriminate between a failed and non-failed firm. Altman and Sabato (2007) have tested the logit model on a sample of US SMEs from various industries. Financial ratios, namely, cash to total assets, earnings before interest, taxes, depreciation, and amortization (EBITDA) to total assets, EBITDA to interest expenses, retained earnings to total assets, and short-term debt to total equity, are found to be the significant indicators of default by SMEs. The overall classification rate of the model was 87.22% in the holdout sample.

Although the logit model does not require the dependent and independent variables to be linearly related, it requires the independent variables to be linearly related to the "logit" function (a ratio of the odds of success to the odds of failure) (Schreiber-Gregory, 2018). Otherwise, the test would underestimate the strength of the relationship and hence may easily reject the relationship (Schreiber-Gregory, 2018).

Alternatively, the ANN model is less stringent (Youn and Gu, 2010b). The key advantage of the ANN model over the other traditional models is its ability to recognize patterns regardless of the functional form (Charitou *et al.*, 2004; Khermkhan and Chancharat, 2013; Kim, 2011). ANNs can examine non-linear relationships and adjust the model adaptively (Sung *et al.*, 1999). However, it functions like a black box, which means the model does not specify if an input variable is positively or negatively affecting the output variable (Montano and Palmer, 2003; Park and Hancer, 2012; Youn and Gu, 2010a). Instead, the ANN model identifies and ranks the input variables based on the degree of significant impact on the output variable.

Despite this shortcoming, ANN is still a good alternative for classifying failed and non-failed firms (Kim, 2018). For example, Odom and Sharda (1990) compared the predictive power of the ANN model to Altman's (1968) MDA model. The five financial ratios used by Altman (1968) were utilized, i.e., sales to total assets, EBIT to total assets, retained earnings to total assets, working capital to total assets, and market value of equity to book value of total liabilities. The ANN model outperformed the MDA model in terms of accuracy rates. It accurately forecast 81.48% of the bankrupt firms in the holdout sample compared to the MDA model that could only accurately predict 59.26% of the bankrupt firms in the holdout sample. Studies have also found that the ANN model performs better than traditional statistical models (Khermkhan and Chancharat, 2013; Kim, 2011; Stroie, 2013). Khermkhan and Chancharat (2013) sought empirical evidence from a sample of Thai SMEs in various sectors. Financial ratios, such as current liabilities to sales, working capital turnover to total assets, return on sales and earnings before interest, and taxes to current liabilities, are used to estimate the models. They reported that the ANN model consistently outperformed the MDA, logit, and probit models, with a correct classification rate of 85% versus 71.6%, 72.8%, and 62.2%, respectively.

Similar results were observed by Stroie (2013) who predicted the failure of Romanian SMEs. The author performed the ANN, logit, and decision tree models using ten accounting ratios that measured firms' profitability, liquidity, leverage, debt coverage, activity, and productivity. Additionally, qualitative information, namely, "management qualification and experience" as well as "maximum delay so far," were also included in the models. The ANN model was found to be marginally better than the logit model, with an accuracy rate of 83.94% in the testing sample, compared to 80.83% and 82.9%, generated by the logit and decision tree models, respectively. Inconsistently, the ANN model has also been found to underperform compared to the MDA and logit models (Altman *et al.*, 1994; Sayeh and Bellier, 2014; Youn and Gu, 2010a). For example, Sayeh and Bellier (2014) found that the logit model (74.8%) yielded a slightly better accuracy rate than the ANN model (71.1%) in the testing sample when they tested their business failure model on a sample of French SMEs. They utilized quantitative and qualitative variables, such as liquid assets to total assets, bank debt to total assets, ROA, tangible assets to total assets, equity to total assets, account receivable turnover, banking relationship, firms' size, firms' age, borrower's gender, and borrower's age, to develop both the logit and ANN models.

As for the hospitality industry, most of the published works have sought evidence from the public listed firms or large firms (Barreda *et al.*, 2017; Fernández-Gámez *et al.*, 2016; Gemar *et al.*, 2019; Gu, 2002; Kim and Gu, 2006; Kim, 2011, 2018; Park and Hancer, 2012; Youn and Gu, 2010a; 2010b; Zhai *et al.*, 2015). Ostensibly, Pacheco (2015) is the only study that has conducted a bankruptcy study of SMEs in the hospitality industry. The study utilized a sample of 25 failed and 460 healthy Portugal SMEs in the restaurant and hotel businesses and reported a prediction accuracy rate of only 68.7% using a logit model. The ratios of debt to total assets and equity to total assets were found to be significant in explaining SMEs' failure.

Moving back to the literature on Malaysia, neither have studies employed the ANN model nor do they predict the failure of SMEs in the hospitality industry. Instead, the only studies are those that model the failure of SMEs in the manufacturing sector (Abdullah *et al.*, 2016a, 2016b). Therefore, this study aims to fill this gap by employing the ANN model and comparing its predictive power against the logit model using a sample of SMEs from the hospitality industry, mainly in the accommodation and restaurant business. From the literature, this study expects the ANN model to generate a better accuracy rate compared to the logit model in predicting the failure of SMEs.

3. Research Methodology

3.1 Sample Specification

A sample of failed and non-failed SMEs in the accommodation and restaurant business were hand collected from the Companies Commission of Malaysia (CCM) database from the period 2000 to 2016. The CCM database provides information such as profiles, balance sheets, and income statements of the companies. Consistent with the National SME Development Council, this study defines SMEs as having annual sales up to RM20 million (SME Corporation, 2013), while failed SMEs are defined as those being wound up by a court order or a creditor's request under Section 218 (1)(e) and (2) of the Companies Act 1965.

The list of SMEs with Malaysia Standard Industrial Classification (MSIC) codes of 5510 (short-term accommodation activities) and 5610 (restaurant activities) was retrieved from the CCM database. The records show active SMEs, SMEs that have been wound up by a court order, a creditor's request, or voluntarily wound up by its members, and companies that had been struck off. Initially, 312 SMEs in the hospitality industry that had been wound up by a court order or by a creditor's request were identified. After data cleaning, only 41 failed SMEs were included in the final sample due to incomplete information. Among these failed SMEs, 10 SMEs were from the accommodation sector and 31 SMEs were from the restaurant business. Subsequently, the failed SMEs were matched with non-failed SMEs on the basis of the same sub-industry and same firm size (± 10%).

This study is designed to predict the failure of SMEs two years before the date of winding up. In many cases, the bankrupt SMEs failed to submit financial statements when their businesses were near bankruptcy. This led to insufficient data for examining the predictors one year before the SMEs declared bankruptcy. Consistent with the earlier studies, 70% of the sample was used for model estimation and the remaining 30% was for the holdout sample to test the predictive power of each model.

3.2 Variables Specification

As there is no unified theory that identifies the appropriate predictors of business failure, the input variables of this study were selected based on the existing literature, subject to the best available data. Four categories of accounting ratios that measured firms' liquidity, leverage, profitability, and efficiency were selected as potential predictors of business failure.

First, liquidity was measured using the *current ratio*, which is the ratio of current assets to current liabilities (Bredart, 2014; Ciampi and Gordini, 2013a; Gunawidjaja and Hermanto, 2010; Wellalage and Locke, 2012). A higher current ratio indicates a better liquidity position for meeting short-term financial commitments and hence a lower business failure risk. Second, the firm's leverage was measured using (1) *debt ratio* and (2) *debt-to-equity*. A higher leverage ratio signifies a higher level of indebtedness and higher financial risk (Abdullah *et al.*, 2016a, 2016b; Pacheco, 2015). Third, profitability measures a firm's ability to generate profits, and it was represented by (1) *return on assets* and (2) *return on sales* (Abdullah *et al.*, 2016b; Altman and Sabato, 2007; Altman *et al.*, 2010; Bredart, 2014; Ciampi and Gordini, 2013b). A higher profitability ratio is negatively related to business failure. Fourth, the firms' efficiency in managing its assets to generate sales was measured by the ratio of sales to total assets or *sales turnover* (Terdpaopong and Mihret, 2011).

The analysis also considered other firm-specific characteristics such as *firm size* and *firm age*, measured by the logarithm form of total assets and firm age in years, respectively (Abdullah *et al.*, 2016a, 2016b; Altman *et al.*, 2010; Lugovskaya, 2010; Wellalage and Locke, 2012). Furthermore, this study examined if governance variables such as ownership concentration, denoted by *OwnC* (Ciampi, 2015), *board size*, and *board gender diversity* (Abdullah *et al.*, 2016b; Ciampi, 2015) are significant in explaining the failure of SMEs. Ownership concentration was represented by a dummy variable that took the value of one if a shareholder held more than 50% of the firm's outstanding shares and zero otherwise (Ciampi, 2015). Board size is the number of directors on boards, while board gender diversity refers to the presence of female directors on boards. It was a dummy variable that took the value of one if at least a woman director was appointed to the boardrooms and zero otherwise.

3.3 Methodology

The logit model incorporates non-linear effects, and the coefficients are estimated based on the cumulative logistic function to predict failure. The coefficients are estimated using the maximum likelihood method that maximizes the probability of classifying the observed data into the appropriate category. A cut-off value of 0.50 was employed to differentiate a failed and non-failed SME. An SME was classified as having failed if the P-value was equal to or greater than 0.50, and it was classified as a non-failed SME if the P-value was less than 0.50. The logit regression functional form is specified below:

$$P = \frac{1}{1 + e^{-(\alpha+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} = \frac{1}{1 + e^{-y}} \tag{1}$$

where P represents the probability of failure, β_n represents the model parameter estimates, and X_n represents the independent variables. The logit model is estimated using a forward stepwise method to identify the most significant predictors that could predict failed and non-failed SMEs.

The second model is the ANN that consists of processing elements or nodes that are connected to a network by an associated value known as "*weight*" (Youn and Gu, 2010a). The common ANN model is the multilayer perceptron (MLP) network (Ciampi and Gordini, 2013a) that is made up of (1) an input layer, (2) an output layer, and (3) one or more hidden layers between the input and output layers. Referring to Zhang *et al.* (1999), this study developed a three-layer MLP network, as shown below.

$$y = f_2(w_2 f_1(w_1 x)) \tag{2}$$

where y is the output, x is the independent variable as the input layer, $w1$ and $w2$ are the matrices of the linking weights from input to the hidden layer and from hidden to the output layer, and $f1$ and $f2$ are the transfer functions for the hidden and output node, respectively. The selection of the transfer function depends on the nature of the output. As the network output lies between 0 and 1, which represents the probability of failure, an appropriate transfer function is the logistic or sigmoid transfer function $f_1(x) = f_2(x) = (1 + e^{-x})^{-1}$. For a binary classification problem, only one output node is required to represent the group membership. SMEs with an output value greater than 0.50 were classified as having failed; otherwise, they were classified as non-failed SMEs.

4. Results and Discussion

4.1 Descriptive Statistics

Table 1 presents the descriptive statistics and the mean difference estimates of the independent variables of the failed and non-failed SMEs. Failed SMEs were significantly less liquid, highly leveraged, less profitable, less efficient, and younger. In terms of the governance variables, the non-failed SMEs were shown to have significantly more board members in the boardrooms compared to the failed SMEs. Both the ownership concentration and board gender diversity were categorical variables. Hence, Table 2 presents the results of the Chi-square test to determine the relationship between the two categorical variables. The results show that the Pearson chi-square test statistic for ownership concentration and board gender diversity is 0.198 and 1.845, respectively. As the p-value is larger than 0.05 for both predictors, the univariate test suggests that there is an insignificant relationship between an SME's failure and ownership concentration as well as an SME's failure and board gender diversity.

Table 1: Summary statistics and mean difference test

Variables	Failed		Non-Failed		Failed–Non-Failed	VIF
	Mean	SD	Mean	SD		
Current ratio	0.5074	0.3805	3.0879	4.0142	-2.5805***	1.222
Debt ratio	1.5888	1.4026	0.4690	0.2975	1.1198***	3.125
Debt-to-equity	-1.9189	6.0858	1.2544	1.6916	-3.1733***	1.269
Return on assets	-0.4710	1.3516	0.1299	0.1113	-0.6009***	2.324
Return on sales	1.8798	2.1189	2.9219	1.8668	-1.0421***	1.462
Sales turnover	-0.4929	1.0930	0.0704	0.1045	-0.5633**	1.393
Firm size	13.4816	1.2632	13.4763	1.2484	0.0053	1.496
Firm age	2.0398	0.7811	2.2937	0.5546	-0.2538*	1.340
Board size	2.6585	0.8547	3.1463	1.1524	-0.4878**	1.239

Notes: *, **and *** significant at 10%, 5% and 1% levels, respectively.

Table 2: Cross tabulation and Chi-Square tests

Variable	Failed	Non-Failed	Total	Pearson Chi-Square	Asymptotic Sig. (2-sided)	VIF
Ownership Concentration				0.198	0.656	1.264
Less or equal 50%	24	22	46			
More than 50%	17	19	36			
Total	41	41	82			
Board Gender Diversity				1.845	0.174	1.269
No woman director on the board	13	19	32			
At least one woman director on the board	28	22	50			
Total	41	41	82			

Table 3 presents the Pearson correlation matrix. The correlations among the variables were relatively low except for *return on assets* and *debt ratio*. Although regression estimates are still unbiased, with the presence of a multicollinearity problem, the estimates are no longer efficient. Thus, to double-check, a variance inflation factor (VIF) test was conducted to confirm that the dataset did not suffer from the problem of multicollinearity. The last column of Table 1 shows the VIF values of each variable, suggesting that the dataset did not suffer from a serious multicollinearity problem.

4.2 Logistic Regression Model

This section discusses the results of the stepwise logistic model. Two predictors were found to be significantly related to the failure of SMEs (refer to Table 4). The insignificant *p-value* of the Hosmer-Lemeshow test suggests that the data fit the model well. The negative coefficient of *return on assets* indicates that SMEs with lower profitability are more likely to fail (Abdullah *et al.*, 2016a, 2016b; Altman *et al.*, 2010; Bredart, 2014; Ciampi and Gordini, 2013b). These SMEs generate insufficient net income from every dollar of assets. Consequently, these SMEs face challenges in meeting the working capital and financial commitments that would ultimately drive the SMEs to bankruptcy. Generating adequate income is critical, especially for SMEs in a capital-intensive industry like those in the hospitality industry (Kim, 2018).

Table 4: Stepwise logistic regression model

Variable	Coefficient	Change in -2 Log Likelihood
Return on assets	-12.422	33.408 (0.000)***
Board size	-1.169	5.768 (0.016)**
Constant	3.441 (0.035)	
Hosmer and Lemeshow test	5.158 (0.741)	

The second significant variable is the board size. The negative coefficient of *board size* reveals that SMEs with larger boardrooms are less likely to fail. This finding agrees with those of the studies by Abdullah *et al.* (2016b) and Keasey and Watson (1987). Generally, a large board is expected to provide better support to the business in terms of experience, expertise, and connections that are important for an SME to operate effectively and continuously in a dynamic environment. Furthermore, a firm is able to grow its customer base and look for better financing opportunities only if the directors have a good network of external parties, as funding is not easily accessible for small enterprises (Audretsch and Mahmood, 1995).

Table 3: Correlation matrix

Variables	Current ratio	Debt ratio	Debt-to-equity	Return on assets	Return on sales	Sales turnover	Firm size	Firm age	OwnC	Board size	Gender diversity
Current ratio	1.0000										
Debt ratio	-0.3290***	1.0000									
Debt-to-equity	0.0636	-0.2330**	1.0000								
Return on assets	0.1533	-0.7230***	0.1026	1.0000							
Return on sales	0.0351	-0.0682	0.3090***	-0.0602	1.0000						
Sales turnover	0.1825	-0.5080***	0.1369	0.3320***	0.2031	1.0000					
Firm size	0.0658	-0.2025	-0.1439	0.1750	-0.4160***	-0.0506	1.0000				
Firm age	0.1678	-0.0284	-0.2270**	0.1488	-0.2270**	-0.0929	0.2910***	1.0000			
OwnC	0.1557	-0.0234	-0.0009	0.0800	-0.1752	-0.0250	0.2530**	0.3540***	1.0000		
Board size	-0.0613	-0.1507	0.0323	0.1106	-0.0563	0.1108	0.1110	0.0079	-0.1308	1.0000	
Gender diversity	0.0000	-0.0932	-0.1065	0.1599	-0.0287	0.1749	-0.1214	-0.0011	0.0025	0.3360***	1.0000

Notes: *, **and *** significant at 10%, 5% and 1% levels, respectively.

Table 5 shows the accuracy rate of the logit regression model. The model accurately classifies 82.8% of the failed SMEs and 89.3% of the non-failed SMEs in the estimation sample. However, in the holdout sample, the logit model performs better in predicting the failed SMEs, compared to the non-failed SMEs, with an accuracy rate of 91.7% and 69.2%, respectively. Overall, the total classification rate is 86% in the estimation sample, but 80% in the holdout sample.

Table 5: Classification accuracy rates of logit model

Percentage correctly classified	Estimation Sample	Holdout Sample
Failed	82.8%	91.7%
Non-failed	89.3%	69.2%
Overall	86.0%	80.0%

4.3 Artificial Neural Network Model

This section discusses the estimates of the three-layer MLP network of the ANN. To begin with, a sensitivity analysis was used to measure the impact of a change in an input variable on an output variable (Montano and Palmer, 2003). The greater the impact observed on the output variable, the greater the sensitivity of the input variable (Montano and Palmer, 2003). In fact, the sensitivity analysis measures the importance of each input variable that has been used to construct the prediction model. Subsequently, the significant predictors will be included in the multilayer perceptron ANN model to predict the failure of SMEs.

Table 6 reports the sensitivity analysis of the observed predictors. Each of the input variables is ranked based on their degree of importance in predicting the output variable. An input variable was inferred to be a significant predictor if the variable's importance value was greater than 0.10 (Youn and Gu, 2010b). The higher the value, the greater the impact of the predictor (input variable) in predicting the failure of SMEs (output variable). Based on the ANN approach, five input variables were found to be significant predictors of the failure of SMEs. The highest-ranked predictor is the *current ratio*, followed by *return on sales*, *debt-to-equity*, *return on assets*, and *board size*, with the variable importance values being greater than 0.1. Looking at the finance-specific variables, liquidity, profitability, and leverage were significant predictors of business failure.

Table 6: Variable importance values

Independent Variables	Rank	Variable Importance Value	Normalized Importance
Current ratio	1	0.159	100.0%
Return on sales	2	0.150	94.2%
Debt-to-equity	3	0.140	88.2%
Return on assets	4	0.140	88.0%
Board size	5	0.102	62.5%
Debt ratio	6	0.079	49.8%
Sales turnover	7	0.071	44.7%
Firm size	8	0.050	31.5%
OwnC	9	0.049	30.6%
Gender diversity	10	0.037	23.1%
Firm age	11	0.024	15.1%

Notes: Normalized importance is the variable importance value divided by the largest importance value.

To compare the performance between the logit and the ANN models, the next step was to determine the accuracy rate of the MLP ANN classification. The results are depicted in Table 7. The ANN model accurately classified 100% of the failed SMEs and 96.4% of the non-failed SMEs in the estimation sample. This resulted in an overall accuracy rate of 98.2%. As for the holdout sample, the model correctly classified 91.7% of failed SMEs and 92.3% of healthy SMEs, with an overall classification rate of 92.0%. Comparing the accuracy rate reported in Table 5 and Table 7, the MLP ANN model outperformed the logit model in predicting business failure in both the training and holdout samples.

Table 7: Classification of accuracy rates of multilayer perceptron ANN model

Percentage correctly classified	Training Sample	Holdout Sample
Failed	100.0%	91.7%
Non-failed	96.4%	92.3%
Overall	98.2%	92.0%

Even though the ANN model is considered as a "black box" model, it is still possible to infer the positive/negative impacts of the significant predictors on business failure, backed by relevant theoretical arguments and empirical evidence. Liquidity, for instance, is the highest-ranked predictor, indicating that liquidity management is an important function for long-term business survival (Mohanty and Mehrotra, 2018). In other words, SMEs with a higher liquidity risk are more likely to fail (Bredart, 2014; Ciampi and Gordini, 2013a; Lugovskaya, 2010). This is because less liquid SMEs have limited cash flows for meeting working capital requirements and for fulfilling debt obligations. Therefore, liquidity-constrained SMEs have a higher probability of failure compared to liquid firms.

In terms of profitability, profitable SMEs are more sustainable compared to SMEs that operate at a loss. SMEs with negative earnings face the challenges of remaining competitive in the market. These SMEs fail to generate enough revenues to cover the costs, or they may do poor cost management. In short, loss-making SMEs have a higher probability to fail (Gu, 2002; Khermkhan and Chancharat, 2013; Lugovskaya, 2010; Stroie, 2013). The third important predictor is a firm's leverage. The adverse effect of leverage on the failure of SMEs can be explained using the trade-off theory. The higher the leverage, the higher the financial risk is. SMEs may end up having financial distress and becoming bankrupt due to poor debt management, which is in line with the findings of Kim (2011). Kim (2011) finds that Korean hotels rely heavily on debt. The proportion of debt is found to be larger than the owners' equity. Financing costs are much higher for these SMEs; this also compresses the firms' profit margin. Hence, debt-burdened hotels, particularly those with poor profitability, are potential candidates for bankruptcy.

According to the resource dependence theory, the board of directors is a mechanism that manages external resources for firms (Pfeffer and Salancik, 1978). It is common for firms that rely more on external resources to appoint the "right" person to sit in the boardrooms. In general, SMEs tend to have a smaller board size, and the board of directors is commonly their family members, who may not have the required external networks or access to external resources for the benefit of their firms. Therefore, having additional directors can further contribute to the success of the firms. This argument is consistent with the reported univariate test, whose results are shown in Table 1, where healthy SMEs were found to have more directors on board, compared to the failed SMEs. In addition, the stepwise logit model also suggests a significant relationship between board size and the failure of SMEs. Therefore, this study argues that in addition to financial predictors, a firm's governance is also key to business survival.

5. Conclusion

This study is designed to answer two research objectives. First, it examines the power of the logit and MLP ANN model in predicting the failure of SMEs in the hospitality industry. Second, it identifies predictors that are significant in determining business failure. The results show that the MLP ANN model has higher prediction accuracy rates for both the estimation and holdout samples than the logit model. In terms of the failure predictors, both models identify a firm's profitability and board size as important determinants of failure. Additionally, the ANN model also identifies liquidity and leverage as predictors of business failure. Therefore, based on the empirical findings, this study argues that the ANN prediction model should be utilized to classify failed and non-failed SMEs because it provides a more accurate and reliable assessment of a firm's financial status.

The findings of this study are subject to several limitations. The analyses were performed on a relatively small sample, and further research should be conducted on a larger sample. Furthermore, the analyses of data two years prior to failure was not sufficient to make strategic plans for recovery. An early warning sign of failure can provide hospitality companies with the time to restructure the companies or take necessary actions to improve the company's profitability, liquidity, and leverage for business survival. Hence, a longer observation period should be taken, such as five years prior to failure to draw more practical implications. In addition, hospitality companies are vulnerable to social and economic changes, as a larger part of their revenue depends on disposable income. Hence, it is worthwhile to consider the external aspects of a company when analyzing failure in the hospitality industry.

References

Abdullah, N. A. H., Ahmad, A. H., Zainudin, N., & Md Rus, R. (2016a). Modelling small and medium-sized enterprises failure in Malaysia. *International Journal of Entrepreneurship and Small Business, 28*(1), 101–116.

Abdullah, N. A. H., Ma'aji, M. M., & Khaw, K. L. H. (2016b). The value of governance variables in predicting financial distress among small and medium-sized enterprises in Malaysia. *Asian Academy of Management Journal of Accounting & Finance, 12*(Suppl. 1), 77–91.

Alaka, H., Oyedele, L.O., Owolabi, H., Kumar, V., Ajayi, S. O, Akinadé, O.O., & Bilal, M. (2018). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Systems with Applications, 94*, 164–184.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609.

Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance, 18*(3), 505–529.

Altman, E. I. & Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the U.S. market. *Abacus*, *43*(3), 332–357.

Altman, E. I., Sabato, G., & Wilson, N. (2010). The value of non-financial information in small and medium-sized enterprise risk management. *Journal of Credit Risk, 6*(2), 1–33.

Audretsch, D., & Mahmood, T. (1995). New firm survival: New results using a hazard function. *The Review of Economics and Statistics, 77*(1), 97–103.

Barreda, A.A., Kageyama, Y., Singh, D., & Zubieta, S. (2017). Hospitality bankruptcy in United States of America: A multiple discriminant analysis-logit model comparison. *Journal of Quality Assurance in Hospitality & Tourism, 18*(1), 86–106.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4,* 71–111.

Bredart, X. (2014). Bankruptcy prediction model using neural networks. *Accounting and Finance Research, 3*(2), 124–128.

Charitou, A., Neophytou, E., & Charalambous, C. (2004). Predicting corporate failure: Empirical evidence for the UK. *European Accounting Review, 13*(3), 465–497.

Ciampi, F. (2015). Corporate governance characteristics and default prediction modeling for small enterprises. An empirical analysis of Italian firms. *Journal of Business Research, 68*(5), 1012–1025.

Ciampi, F., & Gordini, N. (2013a). Small enterprise default prediction modeling through artificial neural networks: An empirical analysis of Italian small enterprises. *Journal of Small Business Management, 51*(1), 23–45.

Ciampi, F. & Gordini, N. (2013b). The potential of corporate governance variables for small enterprise default prediction modeling: Statistical evidence from Italian manufacturing firms: Preliminary findings. *Cambridge Business & Economics Conference Proceedings*, 1-19.

Department of Statistics Malaysia. (2017). Economic census 2016: Profile of small and medium enterprises. Retrieved from https://newss.statistics.gov.my/newss-portalx/ep/epProductCatalogForm.seam?cid=363127

Department of Statistics Malaysia. (2019). Tourism Satellite Account 2018. Retrieved from https://newss.statistics.gov.my/newss-portalx/ep/epFreeDownloadContentSearch.seam?cid=363163

Fernández-Gámez, M. Á., Cisneros-Ruiz, A. J., & Callejón-Gil, A. (2016). Applying a probabilistic neural network to hotel bankruptcy prediction. *Tourism and Management Studies, 12*(1), 40–52.

Gemar, G., Soler, I. P., & Guzman-Parra, V. F. (2019). Predicting bankruptcy in resort hotels: A survival analysis. *International Journal of Contemporary Hospitality Management*, *31*(4), 1546–1566.

Gu, Z. (2002). Analyzing bankruptcy in the restaurant industry: A multiple discriminant model. *International Journal of Hospitality Management, 21*(1), 25–42.

Gunawidjaja, S., and Hermanto, B. (2010). *Default prediction model for SME's: Evidence from Indonesian market using financial ratios* (Research Paper No. 13-04). Retrieved from Social Science Research Network website: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1666703

Jackson, R. H. G., & Wood, A. (2013). The performance of insolvency prediction and credit risk models in the UK: A comparative study. *The British Accounting Review*, *45*(3), 183–202.

Keasey, K., & Watson, R. (1987). Non-financial symptoms and the prediction of small company failure: A test of Argentina's hypotheses. *Journal of Business Finance and Accounting, 14*(3), 335–354.

Khermkhan, J. & Chancharat, S. (2013). Failure forecasting in Thai SMEs: The comparison of the MDA, logit, probit and ANN models. *Proceedings of the Asian Business & Management Conference,* 25-33.

Kim, H., and Gu, Z. (2006). A logistic regression analysis for predicting bankruptcy in the hospitality industry. *The Journal of Hospitality Financial Management*, *14*(1), 17–34.

Kim, S. Y. (2011). Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis. *The Service Industries Journal, 31*(3), 441–468.

Kim, S. Y. (2018). Predicting hospitality financial distress with ensemble models: The case of US hotels, restaurants, and amusement and recreation. *Service Business, 12*(3), 483–503.

Lugovskaya, L. (2010). Predicting default of Russian SMEs on the basis of financial and non-financial variables. *Journal of Financial Services Marketing, 14*(4), 301–313.

Mohanty, B., & Mehrotra, S. (2018). Relationship between liquidity and profitability: An exploratory study of SMEs in India. *Emerging Economy Studies*, *4*(2), 1–13.

Montano, J., & Palmer, A. (2003). Numeric sensitivity analysis applied to feedforward neural networks. *Neural Computing and Applications, 12*(2), 119–125.

Odom, M. D., & Sharda, R. (1990). A neural network model for bankruptcy prediction. *Proceedings of the International Joint Conference on Neural Networks,* San Diego, California, pp. 163–168.

Office for National Statistics (2018). Business demography, UK: 2017, United Kingdom. Retrieved from https://www.ons.gov.uk/businessindustryandtrade/business/activitysizeandlocation/bulletins/businessdemography/2017

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*(1), 109–131.

Pacheco, L. (2015). SMEs probability of default: The case of the hospitality sector. *Tourism and Management Studies, 11*(1), 153–159.

Park, S.-S., & Hancer, M. (2012). A comparative study of logit and artificial neural networks in predicting bankruptcy in the hospitality industry. *Tourism Economics, 18*(2), 311–338.

Pfeffer, J., & Salancik, G. (1978). *The external control of organizations: A resource dependence perspective*. Stanford, CA: Harper & Row.

Sayeh, W., & Bellier, A. (2014). Neural network versus logistic regression: The best accuracy in predicting credit rationing decision. *Proceedings of World Finance Banking Symposium*, 1-28.

Schreiber-Gregory, D. (2018). Logistic and linear regression assumptions: Violation recognition and control. Paper presented at *The SAS Conference Proceedings: Western Users of SAS Software*, Sacramento, California.

Shah, S. H. A., & Ali, A. R. M. (2011). *Entrepreneurship* (2nd ed.). Shah Alam, Malaysia: Oxford Fajar.

SME Corporation. (2013). Guideline for new SME definition. Retrieved from https://www.smecorp.gov.my/images/pdf/Guideline_for_New_SME_Definition_7Jan2014.pdf

SME Corporation. (2019). *SME annual report 2018/19*. Retrieved from https://www.smecorp.gov.my/index.php/en/laporan-tahunan/3911-sme-annual-report-2018-2019

Stroie, L. M. B. (2013). Techniques for customer behaviour prediction: A case study for credit risk assessment. *Proceedings of International Conference on New Techniques and Technologies for Statistics*, 685-694. doi: 10.2901/Eurostat.C2013.001.

Sung, T. K., Chang, N., & Lee, G. (1999). Dynamics of modeling in data mining: Interpretive approach to bankruptcy prediction. *Journal of Management Information Systems, 16*(1), 63–85.

Terdpaopong, K., & Mihret, D. G. (2011). Modelling SME credit risk: A Thai empirical evidence. *Small Enterprise Research, 18*(1), 63–79.

Wellalage, N. H., & Locke, S. (2012). Factors affecting the probability of SME bankruptcy: A case study on New Zealand unlisted firms. *Business Journal for Entrepreneurs (Quarterly)*, 1–17.

Youn, H., & Gu, Z. (2010a). Predict US restaurant firm failures: The artificial neural network model versus logistic regression model. *Tourism and Hospitality Research*, *10*(3), 171–187.

Youn, H., & Gu, Z. (2010b). Predicting Korean lodging firm failures: An artificial neural network model along with a logistic regression model. *International Journal of Hospitality Management, 29*(1), 120–127.

Zhai, S-S., Choi, J-G., & Kwansa, F. (2015). A financial ratio-based predicting model for hotel business failure. *Global Business and Finance Review, 20*(1), 71–86.

Zhang, G., Hu, M. Y., Patuwo, B. E., & Indro, D. C. (1999). Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research, 116(*1), 16–32.

Appendix

Table A: Description of variables

Variables	Description
Current ratio	Current assets to current liabilities
Debt ratio	Total liabilities to total assets
Debt-to-equity	Total liabilities to total equity
Return on assets	Net income to total assets
Return on sales	Net income to net sales
Sales turnover	Net sales to total assets
Firm size	Logarithm form of total assets
Firm age	Logarithm form of firm age in years
OwnC	A dummy variable that is equal to 1 if a shareholder holds more than 50% of the firm's outstanding shares, and 0 otherwise.
Board size	Number of directors in a boardroom
Gender diversity	A dummy variable that is equal to 1 if there is at least one woman director in the boardroom, else 0.