INTERNATIONAL JOURNAL OF INDUSTRIAL MANAGEMENT
ISSN: 2289-9286 E-ISSN: 0127-564X
VOLUME 19, NO. 4, 2025, 173 – 183
DOI: https://doi.org/10.15282/ijim.19.4.2025.12193

RESEARCH ARTICLE

Probabilistic Classification of PN17 Companies Based on Financial Indicators: A Logit Model Approach

Basheer Azdi Shahizan[1], Maisarah Mohd Redwan[2], Wardina Humaira' Rostam[3], Norliza Muhamad Yusof[4]*, Muhamad Luqman Sapaini[4]

[1] Level 30, Tower 2, Petronas Twin Towers, Kuala Lumpur City Centre, 50088 Kuala Lumpur, Malaysia
[2] Level 4 Tower A, Dataran Maybank, No 1 Jalan Maarof, 59000 Bangsar, Kuala Lumpur, Malaysia
[3] Accountant General's Department of Malaysia, Ministry of Finance Complex, No. 1, Persiaran Perdana, Precinct 2, 62594 Putrajaya, Malaysia
[4] Fakulti Sains Komputer dan Matematik, Universiti Teknologi MARA, Cawangan Negeri Sembilan, Kampus Seremban, Malaysia

ABSTRACT – PN17 classification is synonymous with companies that fail to comply with the Bursa Malaysia laws, thus serving as an early warning for potential financial issues. Accurately determining the financial status of PN17 companies often requires an extensive review of the reports and procedures. This study employed a logit model to identify the factors contributing to PN17 classification with a combination of financial distress indicators: stock volatility, leverage, liquidity, profitability, and probability of default (PD). A sample of financial data from 46 companies listed on Bursa Malaysia, comprising both PN17 and non-PN17 companies, from 2017 to 2022 was analysed using logistic regression. The results reveal that stock volatility, liquidity, and leverage are statistically significant to PN17 classification at a 95% confidence interval. The logit function obtained from the logistic regression analysis was able to classify PN17 and non-PN17 companies with 85.1% accuracy. Nonetheless, the accuracy of classifying PN17 (31.9%) is lower compared to non-PN17 (96.4%). This may be due to class imbalance bias (PN17 and non-PN17) and the negative relationship found in leverage, which is contradicted by the financial theory where higher leverage means higher risk. The ROC-AUC analysis supports the struggle in classifying PN17 due to the counterintuitive relationship in leverage. This suggests a complex relationship in leverage that requires further investigation into its linearity, bimodality, and relations with other predictors. Overall, PN17 classification can be improved with stock volatility as the strongest predictor, followed by liquidity and leverage. Meanwhile, profitability and PD were found to be insignificant to PN17 classification. These findings highlight the gaps in the existing literature regarding the predictive power of the logit model when incorporating additional financial indicators (stock volatility and PD) and understanding the relationships of the financial factors in classifying the minority PN17 class. Regulators might consider giving more weight to stock volatility for identifying financially distressed in PN17 companies. Future research may consider data augmentation methods and other model approaches to address the class imbalance bias and counterintuitive relationship in leverage in order to produce a more robust model.

ARTICLE HISTORY

Received	:	05-03-2025
Revised	:	30-06-2025
Accepted	:	15-07-2025
Published	:	05-12-2025

KEYWORDS
Classification
PN17
Financial ratios
Logit model
Regression
Default probabilities

1. INTRODUCTION

In the domain of business financing, identifying the financial health of publicly listed companies is crucial to avoid bad investment decisions. Publicly listed companies are defined under the Malaysian Companies Act 2016, whose shares are publicly listed and traded on Bursa Malaysia, and are subject to its regulatory guidelines (Kiflee & Khan, 2021) Misjudging the financial health of the companies can lead to severe outcomes, including financial losses, poor risk management, delayed responses to financial issues, adverse impacts on portfolio performance, challenges to reputation and credibility, regulatory compliance issues, and market distortions (Ashrofie Hanafi & Ahmed Shahimi, 2020).

The classification of company performance is a critical aspect of financial analysis and strategic decision-making, providing stakeholders with valuable insights into the health and sustainability of a business (Lizares & Bautista, 2021). Recent research by Muparuri and Gumbo (2022) underscores the significance of binary classification in predicting financial distress for companies. Binary classification involves categorising observations into distressed and non-distressed classes, with PN17 classification being a relevant example.

PN17 classification, also known as the regulatory guideline PN17 or Practice Note 17/2005, is a monitoring mechanism established by Bursa Malaysia to improve disclosure practices and address the financial difficulties faced by distressed companies (Ismail et al., 2020). According to Ashrofie Hanafi and Ahmed Shahimi (2020), a publicly listed company is classified as PN17 if it fails to fulfil the minimum capital or equity criteria, which entails having less than

***CORRESPONDING AUTHOR | N. Muhamad Yusof** | ✉ norliza3111@uitm.edu.my

25% of the total paid-up capital, or if it is incapable of meeting its debt commitments. Studies on PN17 companies are becoming increasingly important to accurately assess whether a company is operating consistently at a level deemed satisfactory. The classification of companies as PN17 holds particular significance in the context of financial market regulation and investor protection (Manaf et al., 2021). It serves as an early warning system, signalling financial distress or rules violations that prompt regulatory intervention (Hashim, 2021).

Understanding the financial factors leading to PN17 classification is crucial for policymakers, investors, and stakeholders to formulate effective strategies for financial stability and sustainable economic growth. However, there is limited comprehensive research concerning the specific financial indicators for companies falling under the PN17 category, especially when incorporating factors like stock volatility and probability of default (PD) within a predictive model. Accordingly, the objectives of this study are:

i. To identify the most significant financial indicators (leverage, liquidity, profitability, stock volatility, and probability of default) related to the PN17 classification of companies listed on Bursa Malaysia using the logit model.
ii. To determine the accuracy of the logit model in distinguishing publicly listed companies as PN17 or non-PN17 based on the selected financial indicators.

This study aims to contribute to the existing literature (Ashrofie Hanafi & Ahmed Shahimi, 2020; Azhar et al., 2021; Manaf et al., 2021) by providing a more detailed understanding of the contribution and relationship of the financial factors influencing the classification of PN17 companies. It also provides the predictive power of the logit model in classifying PN17 companies with stock volatility and PD as additional variables. By identifying key financial indicators, this research offers valuable insights for regulators, companies' management, and investors.

2. LITERATURE REVIEW

Financial distress prediction has long been a critical area in financial analysis. The concept of binary classification, which categorises observations into distressed and non-distressed classes, is widely applied in this field (Muparuri & Gumbo, 2022). This approach allows for clear distinctions vital for stakeholders' understanding of a business's health and sustainability (Lizares & Bautista, 2021). Bursa Malaysia's PN17 framework serves as a specific regulatory application of such binary classification to address financial difficulties and improve disclosure practices among distressed companies (Ismail et al., 2020).

PN17 classification is closely linked to specific financial indicators, which are defined as quantitative measurements employed to evaluate the financial well-being, stability, and effectiveness of a company. These indicators serve to identify dangers, risks, and potential imbalances that may exist within the financial system (Hrybinenko et al., 2020) and can encompass both macroeconomic variables and corporation-specific aspects (Zhang et al., 2022). An example of macroeconomic variables is interest rates (Suseno, 2020), while company-specific aspects are leverage, liquidity, and profitability (Haryono et al., 2024; Yuliyanti et al., 2022).

The logit model is chosen as the analytical tool due to its effectiveness in binary classification problems (Zou et al., 2019; Ashrofie Hanafi & Ahmed Shahimi, 2020; Supsermpol et al., 2023). The model is known for its ability to handle categorical outcomes and aligns well with the binary nature of PN17 classification. Additionally, logistic regression highlights the importance of variables in identifying key factors contributing to financial distress in binary classification models. This is evidenced by Muparuri and Gumbo (2022) who used the logit model and metrics like accuracy, sensitivity, specificity, and the ROC curve to categorise companies' financial distress.

Muparuri and Gumbo (2022) also found that the logit model outperformed the artificial neural network model. Other past studies (Mselmi et al., 2017; Christopoulos et al., 2019; Hanafi et al., 2021) have conducted a similar investigation to understand how financial ratios predict financial distress in various contexts using the logit model. The logistic regression model was also reported to often outperform the classification trees and K-nearest neighbour (KNN) approaches of classifying in the medical field (Houfani et al., 2020; Mahi & Zaman, 2023; Joo et al., 2024). The area under the receiver operating characteristic (AUC) for logistic regression was the highest and closer to 1, indicating its robustness in diagnosing gastric mesenchymal tumours (Joo et al., 2024). Logistic regression also performed well in diagnosing diabetes compared to KNN (Mahi & Zaman, 2023). However, the performance of the logit model in predicting financial distress remains debatable.

Previous research by Ashrofie Hanafi and Ahmed Shahimi (2020) used the logit regression analysis to investigate the key factors of PN17, particularly focusing on the financial indicators of leverage (debt ratio and debt-to-equity ratio), efficiency (total assets turnover**),** and profitability (earnings before interest and tax). They found that leverage ratios were positively significant to the financial distress of PN17 companies, while efficiency and profitability ratios were negatively significant. The relationship between financial indicators and financial distress is further supported by Azhar et al. (2021), who reported that leverage was the strongest indicator in predicting corporate failure based on Altman's Z-Score, followed by profitability, liquidity, solvency, and efficiency ratios. Manaf et al. (2021) also found a positive relationship for leverage but a negative and insignificant relationship for profitability. Their study also analysed additional financial indicators, such as liquidity and company size, using Pooled Ordinary Least Squares and Random Effect Model and reported a negative relationship for liquidity, which is consistent with Azhar et al. (2021), and a positive relationship for company size, with both indicators relating to financial distress.

Overall, leverage consistently emerges as the strongest factor of financial distress in PN17 companies. Other indicators, such as liquidity, profitability, solvency, and efficiency, also play a crucial role in identifying PN17 companies (Ashrofie Hanafi & Ahmed Shahimi, 2020; Azhar et al., 2021; Manaf et al., 2021). While previous research has explored the relationships between financial indicators and financial distress in PN17 companies, several inconsistencies remain. For instance, leverage, represented by the debt-to-equity ratio, is said to be insignificant and negatively related to financial distress (Waqas & Md-Rus, 2018), thus contradicting a number of findings (Ashrofie Hanafi & Ahmed Shahimi, 2020; Azhar et al., 2021; Manaf et al., 2021).

Furthermore, macroeconomic factors, such as stock volatility, are increasingly recognised to have an impact on the financial health of PN17 companies (Guo, 2023). Including stock volatility with other financial indicators in logistic regression has received less attention, although it may provide a comprehensive model for PN17 classification. At the same time, PD has been highlighted as another crucial indicator for assessing financial risk, and its relationship with financial ratios is often emphasised in previous studies (Bajwa & Rashid, 2021; Parrado-Martínez et al., 2019). PD measures the likelihood of a company defaulting on its debt obligations. Bajwa and Rashid (2021) highlighted the significant impact of financial ratios on PD, while Parrado-Martínez et al. (2019) advocate its capability to identify various factors affecting a company's risk level. Hence, having an inclusive model with PD and stock volatility as part of the factors in PN17 classification is worth considering.

Although the logit model is commonly used in financial distress prediction, its predictive accuracy, specifically for PN17 classification within the Malaysian context, is less explored and requires further study. This is evidenced in the studies by Azhar et al. (2021) and Manaf et al. (2021), which utilised PN17 companies using other models rather than the dedicated logit model. Baha et al. (2023) attempted to bridge the gap by applying the logit model to evaluate companies' financial distress within the specific context of PN17 classification, particularly when including stock volatility and PD as additional financial indicators. Nevertheless, the relationships and contributions of these financial indicators have been inconsistent in some findings (Waqas & Md-Rus, 2018; Ashrofie Hanafi & Ahmed Shahimi, 2020; Azhar et al., 2021; Manaf et al., 2021). Therefore, this study aims to fill these gaps by combining stock volatility and PD with other proven financial indicators (leverage, liquidity, profitability) into a logit model for PN17 classification. Such an application will determine the predictive power of the logit model and its relationships and significant contributors to PN17 classification.

3. LOGIT MODEL

The logit model describes the relationship between a categorical outcome variable with two or more categories and one or more predictor variables. The logit probability model is expressed as below (Huriyah et al., 2023):

$$logit(p) = ln\left[\frac{p}{1-p}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \quad (1)$$

where:

$p(x)$: probability of occurrence

β_0: the coefficient of a constant term

$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$: the coefficients of the predictor variables x_1, x_2, x_3, x_4, x_5

This logit model is utilised to estimate the probability of PN17 classification, with the dependent variable constrained between 0 (non-PN17) and 1(PN17). The predictor variables, x_1, x_2, x_3, x_4, x_5 are defined as leverage, liquidity, profitability, stock volatility, and PD, respectively.

Null Hypotheses (H0):

There is no significant relationship between financial indicators (leverage, liquidity, profitability, stock volatility, and PD) and the probability of PN17 classification.

Alternative Hypotheses (H1):

There is a significant relationship between financial indicators (leverage, liquidity, profitability, stock volatility, and PD) and the probability of PN17 classification.

4. METHODS AND MATERIALS

4.1 Data Setting

A sample of data from 46 publicly listed Malaysian companies was collected for a duration of six years (2017 to 2022). The data were selected due to the financial data availability and reliability of these companies, thus ensuring the accuracy and consistency of the data used in the analysis. The dataset comprised 24 companies listed as PN17 in 2023 and 22 companies from the main market, representing the non-PN17 companies. Companies under the main market category were used to discriminate against the PN17 companies, since listing on the main market is often associated with larger, more mature companies that meet higher standards of financial performance and governance. All data were obtained from the Eikon DataStream database, involving daily stock prices and financial statements (debt-to-equity ratio, current ratio, gross profit margin, market value, and short- and long-term debts).

The KMV-Merton model was used to estimate PD, which was one of the independent variables in the logit model. The process began by computing the daily book value of debt as the sum of half of the long- and short-term debts. Next, we determined the daily market value of assets by adding the book value of debt to the market value. Daily asset returns were computed by continuously compounding the asset returns, expressed as the natural logarithm of the current asset divided by the previous asset. From these returns, we derived the annual expected asset return, which was calculated as the average of the daily asset returns. On the other hand, annual volatility was calculated by finding the standard deviation of the daily asset returns generated and multiplying it by the square root of 252, which is a common approximation of the number of trading days in a year. Finally, we substituted all values into the following equation derived from the KMV-Merton model to calculate the annual PD (Jumbe & Gor, 2023):

$$PD = N\left(-\frac{ln\left(\frac{A}{D}\right) + (\mu - 0.5\sigma^2)t}{\sigma\sqrt{t}}\right) \tag{2}$$

where:
PD: probability of default
A: market value of assets
D: book value of debt
μ: expected asset returns
σ: asset volatility
t: time

Before conducting the logit regression analysis, descriptive statistical analysis was performed on the independent variables encompassing stock volatility, liquidity (current ratio), profitability (gross profit margin), leverage (debt-to-equity ratio), and the PD estimated from the KMV-Merton model, as shown in Table 1.

Table 1. Descriptive statistics for the independent variables of the logit model

	Leverage	Liquidity	Profitability	Stock Volatility	PD
N	276	270	272	274	276
Min	-1189.07	0	-572.41	0	0
Max	4542.63	19.98	946.66	8.3784	1
Mean	123.1501	1.7072	11.1463	0.7472	0.5459
Std. Dev	396.8349	2.5587	83.3864	0.8262	0.2473
Skewness	5.5866	3.7470	2.8077	5.8829	-0.5176
Kurtosis	58.0577	18.8930	71.6057	49.1366	0.4601

Since the amount of missing data was small and likely missing by chance, some rows with missing data were deleted to obtain the same number of samples (N). Table 1 shows that the data samples contain high skewness and kurtosis with wide ranges between maximum and minimum values, especially in leverage and profitability. A skewness above 1 is considered a large skew and shows asymmetrical distribution in the dataset, while high kurtosis shows the presence of potential outliers where extreme values are more frequent than normal distribution. Therefore, log transformation by adding a constant, log (variable + 1), is necessary to avoid log (0) and to reduce the skewness and kurtosis in the data samples of the independent variables. Log transformation was done to all independent variables, excluding PD, to obtain stable data at once, thus improving the performance and reliability of the logistic regression analysis. Collinearity diagnostic was also performed using indicators such as the Variance Inflation Factor (VIF) and tolerance that were obtained through linear regression.

4.2 Logistic Regression Analysis

The Statistical Package for the Social Sciences (SPSS) software was used to conduct logistic regression analysis on the data. The companies listed in the PN17 category were defined as “1” while those delisted from PN17 or listed under the main market category were defined as “0”. These binary variables acted as the dependent variable of the logit model. Meanwhile, the independent variables encompassed stock volatility, liquidity (current ratio), profitability (gross profit margin), leverage (debt-to-equity ratio), and the PD of companies, which was estimated using the KMV-Merton model.

The logit model’s overall fit was tested using the Omnibus Tests of Model Coefficients, -2 Log likelihood, Cox & Snell R Square, and Nagelkerke R Square. These statistical tests provided insights into the collective significance of the independent variables for the classification of a company as PN17. Furthermore, the model’s discriminatory power was evaluated through the Receiver Operating Characteristic (ROC) Curve and Area Under Curve (AUC), which were fundamental tools for assessing the model's ability to distinguish between PN17 and non-PN17 companies.

5. RESULTS AND DISCUSSION

This study aims to analyse the relationship between the selected financial indicators and PN17 classification using logistic regression analysis, as well as the performance of the logit model. The results are presented in this section, with Table 2 highlighting the key numerical insights and characteristics of the independent variables dataset (financial indicators) utilised in the logistic regression analysis. Meanwhile, Figure 1 shows the distribution of the dependent variable for the logit model, revealing the number of PN17 and non-PN17 cases within the 46 companies sampled from 2017 to 2022.

Table 2. Descriptive statistics for the log-transformed independent variables of the logit model

	Leverage	Liquidity	Profitability	Stock Volatility	PD
N	268	268	268	268	268
Min	0	0	0	0	0
Max	9.11	3.71	7.55	2.88	1
Mean	3.7106	1.1038	2.7274	0.8149	0.5515
Std. Dev	2.5213	0.8257	1.7371	0.4178	0.2451
Skewness	-0.3191	0.532	-0.578	1.224	-0.515
Kurtosis	-1.226	-0.075	-0.931	3.358	0.524

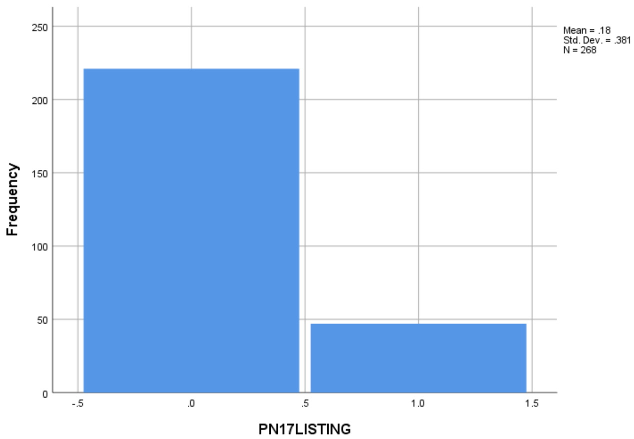

Figure 1. Histogram for the dependent variable of the logit model

Table 2 presents the descriptive statistics for five independent variables of the logit model after log-transformation with a consistent number of 268 cases. The final sample size of 268 cases represents annual observations of leverage, liquidity, profitability, stock volatility, and PD after handling the missing data from 46 publicly listed companies observed over six years (2017-2022). Four of them have a minimum (min) value of 0, namely leverage, liquidity, profitability, and stock volatility, as it suits the log-transformed nature. Meanwhile, PD has maximum (max) and minimum (min) values following the nature of the probabilistic (0-1). The mean values for liquidity, profitability, and stock volatility varied from low to average, with moderate for PD and high for leverage. Combined with the standard deviation, all variables have moderate spreads within that range, except for leverage and profitability.

Both leverage and profitability have high standard deviations, representing wide variability around the mean. The negative skewness found in leverage and profitability indicates that the distribution of the data extends to the lower values. In contrast, the distribution of data for liquidity and stock volatility is positively skewed, describing a tail towards higher values. For PD, it is a direct measure of risk and thus, the negative skewness reflects the natural distributions of risk where most companies have a lower PD with a tail towards a higher value of PD, representing a minority of financial distress companies. The negative kurtosis in leverage, liquidity, and profitability shows that the distribution is flatter than a normal distribution, indicating fewer extreme values in the dataset. In contrast, the distributions for stock volatility and PD have sharper peaks, while stock volatility has the highest kurtosis, implying extreme volatility.

As shown in Figure 1, the dataset for the dependent variable distribution is positively skewed, with 47 and 221 cases classified as PN17 and non-PN17, respectively. This resulted in the mean value of 0.18, indicating that 18% of the 268 observations are classified as PN17. The standard deviation closer to zero (0.381) suggests that most cases fall under the category of non-PN17, where the number of non-PN17 cases is roughly four times the PN17 cases.

Table 3. Logistic regression analysis

Variable	B	S.E.	Wald	df	Sig.	Exp(B)
Leverage	-0.314	0.080	15.266	1	0.000	0.731
Liquidity	-0.701	0.272	6.630	1	0.010	0.496
Profitability	0.149	0.117	1.614	1	0.204	1.160
Stock Volatility	2.264	0.527	18.444	1	0.000	9.617
PD	0.440	0.735	0.358	1	0.550	1.552
Constant	-2.633	0.720	13.378	1	0.000	0.072

Table 3 depicts the logistic regression analysis results within the dataset. It provides an overview of the independent variables contributing to the classification of PN17. Stock volatility emerged as the strongest factor contributing to PN17 classification, as evidenced by the high regression coefficient (B = 2.264) and statistically significant effect (p = 0.000). The coefficients of B and exp(B) indicate that a one unit increase in stock volatility will increase the likelihood for a company to be classified as PN17 by a factor of 9.617. Both leverage and liquidity are also statistically significant with negative coefficients (B = -0.314, p = 0.000 and B = -0.701, p = 0.010), indicating that higher leverage and liquidity are associated with a lower possibility of a company being classified as PN17. For each unit increase in log-transformed leverage and liquidity, the odds of the PN17 listing decrease by a factor of 0.731 and 0.496, respectively.

Furthermore, the Wald statistics of stock volatility, leverage, and liquidity are ranked as the top three highest, indicating their reliability and significance as predictors of the logistic regression models. However, the other variables (profitability and PD) are found to be not statistically significant predictors. Together, these statistical measures conclude that a company is likely to be classified as PN17 when it has higher stock volatility, lower leverage, and lower liquidity. Higher volatility and lower liquidity expose a company to greater risk that can lead to financial distress. Intuitively, these relationships are relevant to the theory of finance, but lower leverage brings a higher possibility of the PN17 classification seems counterintuitive. The distribution of data in leverage with high values of mean and standard deviation might be the reason that the negative relationship makes no sense. Yet, this finding was reported by Waqas and Md-Rus (2018) where leverage (debt-to-equity ratio) was negatively related to financial distress. Overall, this information forms a logit function after removing the insignificant variables, which are defined as follows:

$$y = -0.2.260 - 0.286\, X_1 - 0.580\, X_2 + 2.329\, X_3 \tag{3}$$

where x_1 is leverage, x_2 is liquidity, and x_3 is stock volatility. This logit model (3) assessed its overall performance based on the Omnibus Tests and Model Summary shown in Table 4.

Table 4. Assessing the overall performance of the logit model

Omnibus Tests	Chi-square	66.445
	df	3
	Sig.	0.000
Model Summary	-2 Log likelihood	182.422
	Cox & Snell R Square	0.220
	Nagelkerke R Square	0.363

Table 4 suggests that model (3) provides a significant improvement with leverage, liquidity, and stock volatility as predictors, as indicated by the chi-square 66.45 (df = 3) and p-value (Sig. = 0.000) over a null model. The -2log likelihood = 182.422 describes the overall fit of the model (3). It indicates that the logistic regression model (3) is a relatively better fit than the null model. However, the model tends to underestimate the explained variance in the independent variable as only 22% (Cox & Snell R square = 0.220) of variance is explained by model (3). Moreover, Nagelkerke R square describes a moderate level of explanatory power of 36%. Overall, the Omnibus test suggests that model (3) is statistically valid in classifying PN17 but requires further improvement to explain the variability in the independent variables. The performance analysis of model (3) continues in Table 5, which offers an overview of the observed and predicted PN17 classifications based on the logit model (3).

Table 5. Classification table

Observed	Predicted		
	Non-PN17	PN17	Percentage Correct
Non-PN17	213	8	96.4
PN17	32	15	31.9
Overall Percentage			85.1

Based on Table 5, 213 companies were correctly classified as non-PN17, while 15 companies were correctly classified as PN17. The overall accuracy of the model (3), which is the proportion of correct classifications out of the total number of observations, is 85.1%, which is relatively high. However, the accuracy for the PN17 class is low (31.9%) compared to non-PN17 (96.4%), indicating that the model struggled to identify many of the actual PN17 cases. The low accuracy in classifying PN17 could be related to the counterintuitive leverage that leads to misclassifying the PN17 as non-PN17 companies. Plus, the R square values in Table 3 bring limited predictive accuracy since a moderate amount of variance is explained by the model (3). This can be due to class imbalance bias, which occurs as the data contains an unequal distribution of classes (see Figure 1), leading the model (3) to favour the non-PN17 class and struggle to accurately predict the PN17 class. Accordingly, Table 6 and Figure 2 describe the Area Under the Curve (AUC) and Receiver Operating Characteristic (ROC) curve, respectively. Both are used to evaluate the logit model's (3) ability to discriminate between non-PN17 and PN17.

Table 6. Area under the curve (AUC)

Test Result Variable(s)	Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
Stock Volatility	0.838	0.033	0.000	0.774	0.902
Liquidity	0.613	0.039	0.003	0.537	0.689
PD	0.557	0.056	0.307	0.448	0.666
Profitability	0.498	0.055	0.971	0.391	0.605
Leverage	0.245	0.038	0.000	0.170	0.321

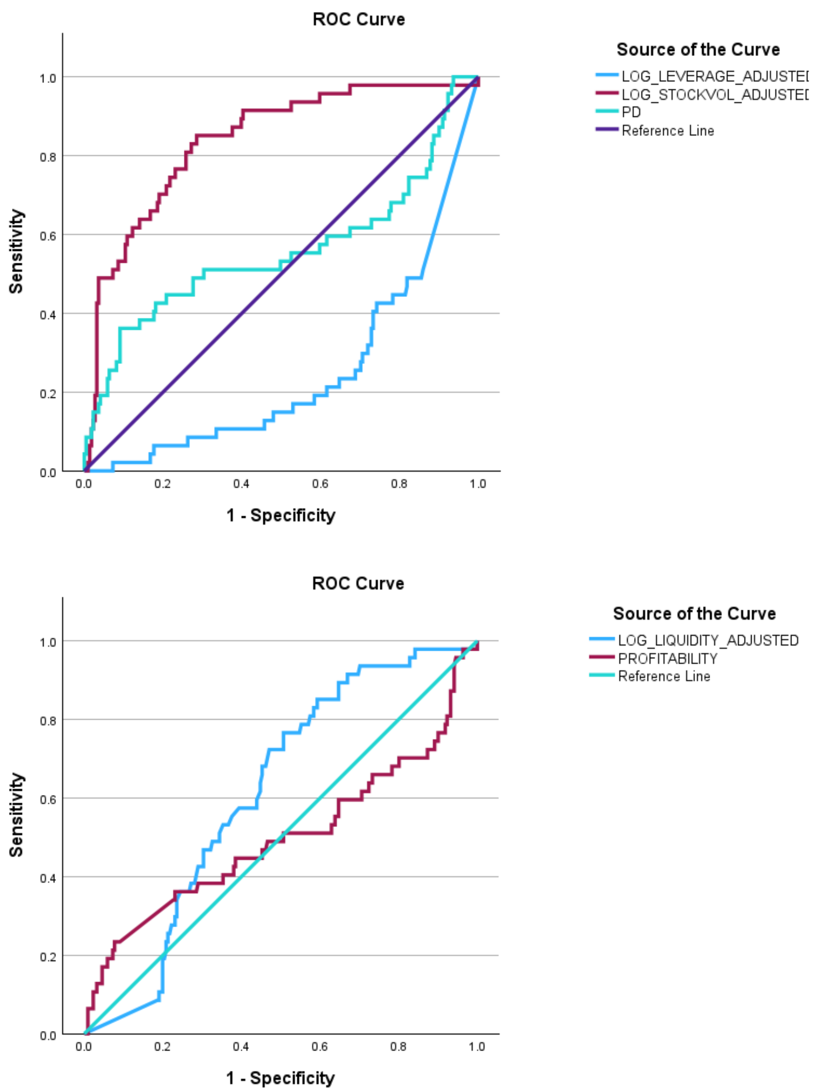

Figure 2. ROC Curve

According to Table 6, stock volatility and liquidity appear to be effective predictors since the AUC values of both variables are more than 0.5 (refer to area values) and significant (refer to asymptotic sig.). Nevertheless, stock volatility appears to be the most effective predictor as it has an AUC close to 1 (0.838) and a narrow asymptotic 95% confidence interval compared to liquidity. The same conclusion can be seen in Figure 2, where the curve for stock volatility (refer to the red curve defined as log stockvol_adjusted) describes a steeper curve above the reference line, approaching the left top, which indicates better sensitivity and specificity. Therefore, stock volatility and liquidity present better individual predictive power than random guessing. On the other hand, PD and profitability remain insignificant. This is consistent with the results in Table 3 regarding the counterintuitive relationship of leverage. The ROC curve of leverage (refer to the blue curve denoted as log_leverage_adjusted) is found below the reference line with an AUC less than 0.5 (Area = 0.245). It indicates that low leverage is more likely to be in the PN17 category, which is the opposite of the financial theory expectation. This can also be the reason why the model (3) struggles to accurately predict the minor class (PN17), aside from the imbalance class bias discussed in Table 5.

Figure 3 provides reasons for the negative relationship found in leverage towards the PN17 classification despite the positive relationship suggested in the ROC curve. The histogram in Figure 3 shows a bimodal distribution, meaning that there are two clusters in the leverage dataset. Cluster 1 presents companies with low leverage at the left tail, while cluster 2 presents companies with high leverage at the right tail. The gaps between the clusters suggest a difference between the two group companies in terms of leverage. Bimodal distribution causes a negative coefficient in the logistic regression analysis (Table 3), which can be attributed to the non-linear relationship between leverage and PN17 status, although log transformation has been made. Moreover, the effects of a large proportion of low-leverage clusters in the non-PN17 category made the relation seem negative, even though the right relationship in the high-leverage cluster is positive. Overall, the PN17 classification can be better with stock volatility and liquidity as predictors. However, considering leverage might limit the logit model's (3) performance.

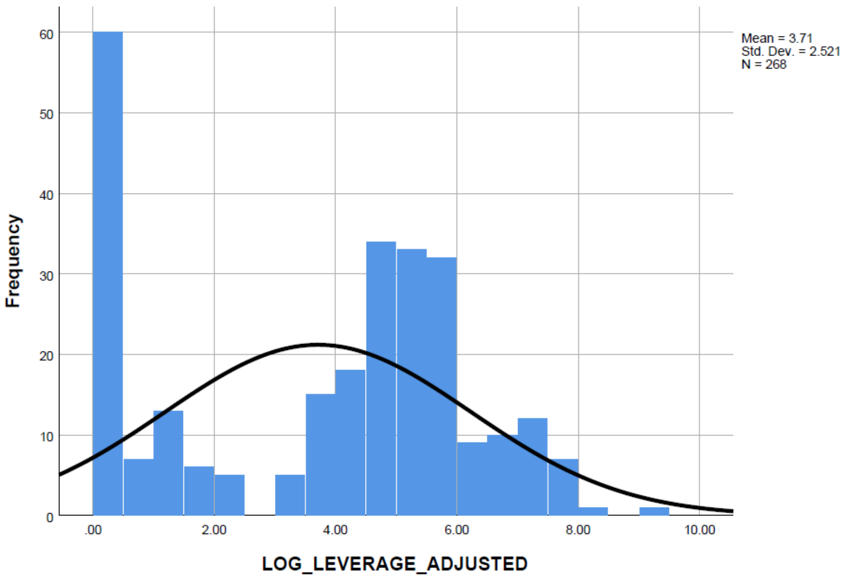

Figure 3. Histogram of log-leverage data distribution

6. CONCLUSION

This study presents the challenges in achieving high accuracy when classifying Malaysian publicly listed companies as PN17 using the logit model due to the minority class within the dataset. There is a need for models that consider class imbalance bias to provide more reliable early warning of PN17 classification. Furthermore, the lack of studies on PN17 classification with specific key indicators (stock volatility and PD) in existing logit models warrants attention to better understand the contributions and relationships of the financial factors to PN17 classification. This study fills the notable gap by delving into the intricacies of classifying selected Malaysian publicly listed companies as PN17 based on pre-determined financial indicators, namely leverage (debt-to-equity ratio), liquidity (current ratio), profitability (gross profit margin), stock volatility, and PD, using the logistic regression model.

The findings show that stock volatility is the best predictor for classifying PN17 companies, followed by liquidity (current ratio). Both variables were found statistically significant at a 95% level in the logistic regression analysis and ROC-AUC. The increment in stock volatility and decrement in liquidity can lead to poor financial health, resulting in a higher probability of PN17 classification. Also, leverage shows consistency as a substantial financial factor in PN17 classification, which agrees with the findings of previous studies (Ashrofie Hanafi & Ahmed Shahimi, 2020; Azhar et al., 2021; Manaf et al., 2021). Yet, the negative relationship of leverage (debt-to-equity ratio) conflicts with the nature of financial theory, as reported by Waqas and Md-Rus (2018). Leverage has an inverse relationship with PN17 classification, where lower leverage is associated with a higher rate of being classified as PN17. This might be due to the complex relationship presented in the bimodal distribution of the leverage dataset. Meanwhile, profitability (gross profit margin) and KMV-Merton's PD are found to be insignificant to PN17 classification.

Furthermore, the logit model improves PN17 classification with stock volatility, liquidity, and leverage as predictors, as evidenced by the high chi-square 66.45 (df = 3) and significant p-value (Sig. = 0.000). Overall, the logit model's accuracy in classifying PN17 (31.9%) and non-PN17 (96.4%) is 85.1%. However, class imbalance bias was detected as the logit model's classification was biased towards the non-PN17 class. This is because the data contain an unequal distribution of classes. Further ROC-AUC analysis revealed that the effects of a substantial proportion of low-leverage clusters in the non-PN17 category may be the reason for the logit model struggling to predict the minor class (PN17) accurately. Therefore, further research should explore factors like data quality and model complexity. To confirm class imbalance bias, it is essential to consider additional metrics and techniques, such as the confusion matrix, precision, recall, F1-score, and class distribution analysis.

The findings of this study offer crucial insights for regulators, particularly Bursa Malaysia, to incorporate stock volatility as part of the primary indicators in developing an early warning system for PN17 classification, particularly among publicly listed companies in Malaysia. Companies' management can also proactively monitor their own financial health through their stock volatility, liquidity, and leverage levels before they escalate to PN17 status.

ACKNOWLEDGEMENTS

This research paper would not have been possible without the support of facilities and resources from UiTM Cawangan Negeri Sembilan, Kampus Seremban. Also, our sincere gratitude goes to the editorial team and reviewers for their contributions, which strengthened the clarity and consistency of this research paper.

FUNDING STATEMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AUTHORS CONTRIBUTION

Basheer Azdi Shahizan: Conducting the research and investigation process, specifically collecting, implementing, analysing, and interpreting the data. Also, writing the initial draft of the article.

Maisarah Mohd Redwan: Conducting the research and investigation process, specifically collecting, implementing, analysing, and interpreting the data. Also, writing the initial draft of the article.

Wardina Humaira' Rostam: Conducting the research and investigation process, specifically collecting, implementing, analysing, and interpreting the data. Also, writing the initial draft of the article.

Norliza Muhamad Yusof: Designing the research work and writing the article.

Muhamad Luqman Sapini: Reviewing and editing the article.

AVAILABILITY OF DATA AND MATERIALS

The datasets used in this research are available and can be provided upon request.

ETHICS STATEMENT

This research complies with ethical standards, ensuring informed consent, confidentiality, and the privacy of all participants.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

GENERATIVE ARTIFICIAL INTELLIGENCE DECLARATIONS

The author(s) declare that no generative AI or AI-assisted technologies were used in the writing of this manuscript. All content, including text, figures, and tables, was created by the author(s).

REFERENCES

Ashrofie Hanafi, A. H., & Ahmed Shahimi, W. R. M. (2020). What drives financial distress for PN4 and PN17 listed firms in Malaysia? Advanced International Journal of Banking, Accounting and Finance, 2(4), 01–09.

Azhar, N. I. H., Lokman, N., Alam, Md. M., & Said, J. (2021). Factors determining Z-score and corporate failure in Malaysian companies. International Journal of Economics and Business Research, 21(3), 370–386.

Baha, R., Lévy, A., & Hasnaoui, A. (2023). Capital structure and default risk of small and medium enterprises: evidence from Algeria. The Journal of Risk Finance, 24(4), 523-536.

Bajwa, A. M. A. & Rashid, M. (2021). Financial ratios: A tool for computing probability of corporate default. Audit and Accounting Review, 1(2), 71–90.

Christopoulos, A. G., Dokas, I. G., Kalantonis, P., & Koukkou, T. (2019). Investigation of financial distress with a dynamic logit based on the linkage between liquidity and profitability status of listed firms. Journal of the Operational Research Society, 70(10), 1817–1829.

Guo, Q. (2023). The relationship between investor sentiment and stock market price. Frontiers in Business, Economics and Management, 9(2), 124–129.

Hanafi, A. H. A., Md-Rus, R., & Mohd, K. N. T. (2021). Predicting financial distress in the Malaysian market: Hazard model versus the logit model. Advanced International Journal of Banking, Accounting and Finance, 3(7), 01–19.

Haryono, C. B. C., Worokinasih, S., & Darmawan, A. (2024). The mediating role of dividend policy on financial performance and firm value: A literature review. KnE Social Sciences, 9(13), 247–257.

Hashim, H. N. M. (2021). Why financially distressed companies PN17 of Bursa Malaysia regularisation plan are delayed? [Master's thesis, Universiti Tun Abdul Razak]. https://kmc.unirazak.edu.my/wp-content/uploads/2022/06/Master-Thesis-Hisham.pdf

Houfani, D., Slatnia, S., Kazar, O., Zerhouni, N., Saouli, H., & Remadna, I. (2020). Breast cancer classification using machine learning techniques: a comparative study. Medical Technologies Journal, 4(2), 535-544.

Huriyah, S. K., Satyahadewi, N., Perdana, H., & Tamtama, R. (2023). Binary logistics regression to predict the opportunity of SNMPTN graduation in statistics study program of Tanjungpura University. Mathline: Jurnal Matematika dan Pendidikan Matematika, 8(1), 17–28.

Hrybinenko, O. Bulatova, O. & Zakharova, O. (2020). Financial indicators in the system of economic security of the world countries. In 11th International Scientific Conference "Business and Management 2020".

Ismail, A. M., Xing, W. X., Bi, R., Sadique, M., Zakaria, N. B., Alam, P., Services, R., Pacific, A., & Bhd, S. (2020). Corporate governance and earnings management: evidence from financially distressed (PN17) companies in Malaysia. Corporate Governance, 11(10), 765–789.

Joo, D. C., Kim, G. H., Lee, M. W., Lee, B. E., Kim, J., & Kim, K. (2024). Artificial intelligence-based diagnosis of gastric mesenchymal tumors using digital endosonography image analysis. Journal of Clinical Medicine, 13(13), 1-12.

Jumbe, G., & Gor, R. (2023). Examining the impact of debt maturity time, expected return and volatility on probability of default in credit risk modelling: The case of Merton and MKMV Models. International Journal of Engineering Science Technologies, 7(1), 29–42.

Kiflee, A. K. R. & Khan, M. N. A. A. (2021). The effect of performance and corporate governance to risk disclosure among listed companies in Malaysia. Asia-Pacific Management Accounting Journal, 16(1), 119-161.

Lizares, R. M., & Bautista, C. C. (2021). Corporate financial distress: The case of publicly listed firms in an emerging market economy. Journal of International Financial Management and Accounting, 32(1), 5–20.

Mahi, H. M. and Zaman, A. S. (2023). A comparative analysis of logistic regression and k-nearest neighbours algorithms in diagnosis of diabetes. GANIT: Journal of Bangladesh Mathematical Society, 43(2), 01-07.

Manaf, S. M. A., Amzah, N., & Salleh, W. A. (2021). An investigation of factors affecting financial distress: Analysis among PN17 companies in Bursa Malaysia. Insight Journal, 8, 182-198.

Mselmi, N., Lahiani, A., & Hamza, T. (2017). Financial distress prediction: The case of French small and medium-sized firms. International Review of Financial Analysis, 50, 67–80.

Muparuri, L., & Gumbo, V. (2022). On logit and artificial neural networks in corporate distress modelling for Zimbabwe listed corporates. Sustainability Analytics and Modeling, 2, 1-11.

Parrado-Martínez, P., Gómez-Fernández-Aguado, P., & Partal-Ureña, A. (2019). Factors influencing the European bank's probability of default: An application of SYMBOL methodology. Journal of International Financial Markets, Institutions and Money, 61(April 2014), 223–240.

Supsermpol, P., Huynh, V. N., Thajchayapong, S., & Chiadamrong, N. (2023). Predicting financial performance for listed companies in Thailand during the transition period: A class-based approach using logistic regression and random forest algorithm. Journal of Open Innovation: Technology, Market, and Complexity, 9(3), 1-16.

Suseno, M. A. (2020). Financial performance, macroeconomic factors and company characteristics in consumer goods company in Indonesia. International Journal of Economics, Business and Accounting Research, 4(03), 296–304.

Waqas, H., & Md-Rus, R. (2018). Predicting financial distress: Importance of accounting and firm-specific market variables for Pakistan's listed firms. Cogent Economics & Finance, 6(1), 1-16.

Yuliyanti, L., Waspada, I., Sari, M., & Nugraha, N. (2022). The effect of liquidity, leverage, and profitability on firm value with firm size as moderating variable BT - Proceedings of the 6th Global Conference on Business, Management, and Entrepreneurship (GCBME 2021). 88–92.

Zhang, X., Zhao, Y., & Yao, X. (2022). Forecasting corporate default risk in China. International Journal of Forecasting, 38(3), 1054–1070.

Zou, X., Hu, Y., Tian, Z., & Shen, K. (2019). Logistic regression model optimization and case analysis. Proceedings of IEEE 7th International Conference on Computer Science and Network Technology, ICCSNT 2019, 135–139.