JURNAL ECONOMINA
e-ISSN: 2963-1181
Volume 5, Number 10, October 2026
Homepage: ejournal.45mataram.ac.id/index.php/economina

Predicting the Risk of MSME Bankruptcy in Kupang City: Integrating the Altman Z-Score, Macroeconomic Impact, Machine Learning, and Explainable AI (SHAP and LIME)

Olaf Tri Wilopo Simanjuntak[1*], Fajri Adha[2], Theddy Baskhara Foenale[3], Septimus Adrianus Mali[4], Nur Muhlis[5], James Miguel Danielo[6], Rahmat Joshua Setia Benu[7], Kayleen RI Bale Doto[8], Ilham Samuel Silahooy[9]

[1*,3,4,5,6,7,8]Business Administration Study Program, Faculty of Social and Political Sciences (FISIP), Nusa Cendana University, Indonesia
[2,9]Computer Science Study Program, Faculty of Science and Engineering (FST), Nusa Cendana University, Indonesia

**Corresponding Author: olaf.simanj untak@staf.undana.ac.id*

Article History

Received: 23-10-2026
Revised: 29-10-2026
Published: 30-10-2026

ABSTRACT

This study aims to assess the financial health and predict the bankruptcy risk of 100 micro, small, and medium enterprises (MSMEs) in Kupang City by integrating the Altman Z″ score, testing the impact of macroeconomic variables, *machine learning modeling* , and *explainable artificial intelligence* (XAI). Data completeness was tested with Little's MCAR test, the impact of macroeconomics (inflation, deposit interest rates, and rupiah exchange rates) on the Z″ score was tested with the Wilcoxon Signed-Rank Test, and four classification algorithms (Gradient Boosting, Random Forest, Logistic Regression, and Naive Bayes) were compared on 11 selected features, then the best model was explained with SHAP and LIME. The results showed that missing data was perfectly random (Sig. 0.587) and all MSMEs experienced a decrease in health scores after macroeconomic factors were included (Z = −8.697; Sig. 0.000). Gradient Boosting provides the best performance in accuracy, precision, *recall* , and F1-score (80%; 0.83; 0.71; 0.77) with an AUC of 0.85. SHAP analysis places D2, total current assets, and D1 indicators as global determinants, while LIME on three MSMEs shows that low D2 and limited total current assets drive the prediction of bankruptcy risk in MSME #2. These findings confirm that assessments based on internal data alone tend to be overly optimistic and that the combination of predictive models with transparent explanations can be the basis of an early warning system for businesses, banks, and local governments.

Keywords: MSME Bankruptcy, Altman Z-Score, Macroeconomics, Gradient Boosting, Explainable AI, Kupang City

INTRODUCTION

Micro, small, and medium enterprises (MSMEs) are a dominant group of business entities in Indonesia and contribute a significant portion to gross domestic product (Tambunan, 2019). However, MSMEs generally operate with limited financial records, resulting in limited

information available to assess their resilience, both for owners, financing institutions, and business managers (Wahlstrøm et al., 2024). This situation makes early detection of financial distress a crucial issue: the failure of an MSME not only eliminates the owner's livelihood but also disrupts the supply chain and employment opportunities in the surrounding area.

Bankruptcy prediction is classically performed using a financial ratio-based scoring model. Altman (1968) introduced the Z-score through multiple discriminant analysis, and the model has been continuously reviewed and adjusted for over five decades (Altman, 2018; Altman et al., 2017). For non-manufacturing companies and emerging markets, a variant of the Z″ is available that does not include the sales-to-assets ratio (Altman, 2005), making it more suitable for small businesses with diverse asset characteristics. In Indonesia, the application of the Z-score to MSMEs is generally descriptive, for example, categorizing businesses into safe, gray, and potentially bankrupt zones (Romaniyah & Alfian, 2024). This approach is static because it only utilizes internal data and does not account for economic environmental pressures, even though Tinoco and Wilson (2013) showed that adding macroeconomic variables enriches the prediction of financial distress.

Advances in *machine learning* have opened up more flexible alternatives. Barboza et al. (2017) reported that machine learning models, particularly ensemble methods, outperformed classical statistical models in bankruptcy prediction. In the MSME segment, Shetty et al. (2022) obtained a global accuracy of 82–83% on Belgian MSME data with three financial ratios, Correa (2023) found that XGBoost outperformed logistic regression on Colombian MSMEs, and Káčer and Alexy (2025) showed that finely tuned machine learning models generally outperform traditional statistical models on Slovakian MSMEs. Kou et al. (2021) also utilized transactional data and two-stage feature selection to predict MSME bankruptcy. The *gradient boosting algorithm* formulated by Friedman (2001) forms the basis of most of these models. However, predictive accuracy alone is insufficient for financial decision-making. Ensemble models are black-box , preventing users from exploring the rationale behind a prediction (Barredo Arrieta et al., 2020).

Explainable artificial intelligence (XAI) addresses these issues. SHAP (Lundberg & Lee, 2017; Lundberg et al., 2020) allocates the contribution of each feature to a prediction based on its Shapley value and can be interpreted globally or locally, while LIME (Ribeiro et al., 2016) explains a single prediction through a simple local surrogate model. Both methods have been applied to bankruptcy and credit risk prediction (Bussmann et al., 2021; Gramegna & Giudici, 2021; Nguyen et al., 2025; Park et al., 2021), and a recent review charts their development in finance (Yeo et al., 2025).

This study selected MSMEs in Kupang City, the capital of East Nusa Tenggara Province, as its object for several reasons. First, Kupang City represents a small-scale regional economy whose business owners are dependent on economic conditions beyond their control. Second, the sample data exhibits high heterogeneity, ranging from stall businesses with total assets of Rp1,000,000 to businesses with total assets of Rp3,000,000,000. Third, cash transaction practices mean that many MSMEs do not have accounts receivable and payable: only 22 out of

100 MSMEs filled in accounts receivable and 55 filled in accounts payable. This data characteristic is rarely captured in studies based on financial statements of registered companies or registry data.

From the description above, there are three research gaps. (1) Z-score studies on Indonesian MSMEs generally stop at score classification and have not statistically tested the impact of macroeconomic variables on them. (2) The *machine learning- based MSME bankruptcy prediction studies* referred to above rely on data from other countries, while XAI studies on bankruptcy generally use company or credit data, not primary data from MSMEs in regional cities. (3) To the author's knowledge, there are not many studies that combine financial health scores, macroeconomic impact tests, algorithm comparisons, and global and local explanations in a single analytical framework. This study offers such an integrated framework with the following objectives: (1) testing missing data patterns with Little's MCAR test; (2) testing whether macroeconomics changes MSME financial health scores; (3) comparing the performance of four classification algorithms in predicting bankruptcy risk; and (4) explaining the best model globally with SHAP and locally with LIME.

RESEARCH METHODS

Design, data, and variables

This study uses a quantitative approach with 100 MSMEs in Kupang City as the unit of analysis. Financial data includes 12 accounts in rupiah, namely cash and cash equivalents, accounts receivable, inventory, total current assets, total assets, total current liabilities, total debt, total equity/capital, annual net sales, annual net profit, retained earnings/accumulated profits, and *earnings before interest and taxes* (EBIT). The dataset for modeling contains 33 numerical features consisting of financial and macroeconomic variables, including indicators D1 and D2. The macroeconomic variables used in the impact test on the health score are inflation, deposit interest rates, and the rupiah exchange rate.

Data preprocessing and missing data testing

Data completeness was checked through *Missing Value Analysis* with *the Expectation-Maximization* (EM) algorithm in SPSS. Little's MCAR test (Little, 1988) tests the null hypothesis that missing data occurs at perfect random. Data are declared to meet MCAR if Sig. > 0.05; if Sig. < 0.05, the missing data pattern is systematic and data deletion can cause bias. Empty cells in accounts receivable and payable are interpreted as zero values because MSMEs with cash transactions do not record these accounts, so that all 100 MSMEs can still calculate the ratio. In the *machine learning modeling stage* , empty values in numeric features are imputed with the mean (*mean imputation*) and features are standardized with *StandardScaler*.

Outlier detection and descriptive statistics

Univariate outliers were identified using two criteria: standardized Z-score, where values outside the range of −3.0 to +3.0 were considered outliers, and the interquartile range (IQR)

method, where values outside Q1 − 1.5 × IQR to Q3 + 1.5 × IQR were considered mild outliers and those outside 3.0 × IQR were considered extreme outliers. Outlier values were not discarded because they reflect heterogeneity in capital and debt capacity among MSMEs. Descriptive statistics are presented in the form of the number of valid data, minimum, maximum, mean, and standard deviation values.

Altman Z" score and Wilcoxon test

Financial health is calculated using the modified Altman Z″ model (Altman, 2005) which contains four ratios: X1 (working capital to total assets), X2 (retained earnings to total assets), X3 (EBIT to total assets), and X4 (equity to total debt). Based on the scores, MSMEs are grouped into four categories: Very Healthy, Healthy, Fairly Healthy/Wary, and Troubled/Critical. The macroeconomic impact is tested by comparing the score incorporating macroeconomic factors (InEx_ZScore) to the internal data-based score (In_ZScore) using *the Wilcoxon Signed-Rank Test* (Wilcoxon, 1945) for two paired samples. The nonparametric test was chosen because the distribution of scores is very skewed (see the results section). The decision was taken at the significance level of $\alpha = 0.05$: Sig. < 0.05 means there is a significant difference between the two scores.

Machine learning modeling

The target variable was formed from ten business stress indicators (E1–E10) in Yes/No format. MSMEs with four or more "Yes" answers were labeled as at risk of bankruptcy (1), the rest as healthy (0), resulting in 57 healthy MSMEs and 43 at risk. Feature selection was performed using *Random Forest* (Breiman, 2001): features whose *importance* exceeded the mean (0.0303) were retained, reducing the initial 33 features to 11. The data were divided into 70% training data and 30% testing data (30 MSMEs) with 42 *random states* . Four algorithms were compared with the default parameters *of scikit-learn* (Pedregosa et al., 2011): Logistic Regression, Random Forest, Gradient Boosting (Friedman, 2001), and Naive Bayes. Performance was measured by accuracy, precision, *recall* , F1-score, *area under the curve* (AUC), and confusion matrix. The best model was selected based on the balance of these metrics.

Explainable AI

The best models were explained using two methods. SHAP was calculated using *TreeExplainer* on the test data; the mean absolute SHAP value (*mean* |*SHAP*|) was used to rank features globally, while the distribution of points in *the summary plot* was used to read the direction of influence. Positive SHAP values pushed predictions toward bankruptcy risk and negative values toward health risk. LIME (*LimeTabularExplainer* , classification mode) explained individual predictions through the top ten features on the first three MSMEs in the test data, designated MSMEs #1, #2, and #3. The numbering indicates the order of samples in

the test data, not the order of model trials. Because features have been standardized, the range of values in LIME explanations is expressed in standard deviation units.

RESULTS AND DISCUSSION

Completeness of data and characteristics of MSMEs

Little's MCAR test yielded *a Chi-Square* of 114.007 with 118 degrees of freedom and a Sig. 0.587. Because these values exceed 0.05, the hypothesis that missing data occur at perfect random is not rejected, so there is no evidence of a systematic missing pattern and the data are suitable for proceeding to the ratio calculation. Two caveats are worth noting. First, the EM algorithm did not converge in 25 iterations, so the test results should be read as indicative, not absolute proof. Second, these results do not contradict the interpretation of zeros in accounts payable and receivable: the MCAR test assesses whether missingness is related to the value of another variable, while the interpretation of zeros is based on the cash-transacting nature of the business. Table 1 shows that accounts receivable were only filled in 22 MSMEs, total current liabilities in 47 MSMEs, and total payables in 55 MSMEs, so only 12 MSMEs filled in all 12 accounts simultaneously (*valid N listwise*). Without zero filling, the listwise analysis would have discarded 88 MSMEs.

Table 1. Descriptive statistics of raw financial data of MSMEs (in rupiah)

Variables	N	Minimum	Maximum	Average	Standard deviation
Cash and cash equivalents	97	5,000	500,000,000	16,893,556.70	62,260,405.26
Accounts receivable	22	12,000	71,000,000	5,476,909.09	15,293,810.16
Supply	98	70,000	1,000,000,000	41,809,612.24	134,683,565.94
Total current assets	100	200,000	1,002,000,000	58,315,090.00	149,862,839.02
Total assets	100	1,000,000	3,000,000,000	177,590,340.00	495,894,666.45
Total current liabilities	47	15,000	96,000,000	7,860,414.89	16,603,622.99
Total debt	55	15,000	500,000,000	29,597,081.82	71,914,859.19
Total equity/capital	99	-48,250,000	2,904,000,000	162,769,641.41	472,871,612.19
Annual net sales	100	1,296,000	4,800,000,000	196,161,960.00	512,652,685.26
Annual net profit	100	200,000	500,000,000	69,383,600.00	88,504,798.86
Retained earnings/accumulated earnings	99	0	300,000,000	15,020,101.01	42,529,967.45
EBIT	100	200,000	500,000,000	71,347,147.08	90,031,096.64

Source: SPSS data processing results.

The standard deviation of all variables exceeds the mean. For example, total assets have a mean of Rp177,590,340.00 with a standard deviation of Rp495,894,666.45. This pattern indicates that the sample combines small-scale micro-enterprises with established businesses, and the mean does not adequately reflect a "typical" MSME. The minimum total equity/capital is negative (−Rp48,250,000.00), indicating an MSME with accumulated losses to the point where its debt exceeds its assets. This heterogeneity is important because it explains why MSME financial ratios easily produce extreme values, as seen in the discussion of the Z'' score below.

Financial health based on Altman Z" score

Table 2 summarizes the distribution of financial health categories. Of the 100 MSMEs, 37 are classified as Very Healthy, 12 as Healthy, 19 as Fairly Healthy/Alert, and 32 as Problematic/Critical. This distribution is polarized: more than a third of MSMEs are in the highest category, while almost a third are in the lowest category, and only 12 MSMEs are in

the Healthy category. Z″ scores range from -52.61 to 1,648.53 with a mean of 115.30, which is well above the median of 58.30, indicating a distribution that is strongly skewed to the right.

Table 2. Distribution of MSME financial health categories based on internal Z″ scores

Category	Number of MSMEs	Percentage	Observed Z″ score range
Very Healthy	37	37%	80.02 – 1,648.53
Healthy	12	12%	63.25 – 78.94
Fairly Healthy/Alert	19	19%	40.56 – 58.33
Problematic/Critical	32	32%	−52.61 – 38.65
Total	100	100%	−52.61 – 1,648.53

Source: Results of calculating the X1–X4 ratio. The range is the minimum and maximum values observed in each category.

These extreme scores should be read with caution. The highest score (1,648.53) for a vegetable-selling MSME was driven almost entirely by X4 of 1,608.95, a ratio of equity to total debt that increased due to its very low debt. This means that some "Very Healthy" scores reflect a near-zero denominator, not truly robust business resilience. Debt-based ratios become unstable in businesses with virtually no debt, and this is one reason why a single score needs to be complemented by a model that studies patterns across multiple variables.

Macroeconomic impact on financial health scores

The Wilcoxon test shows that all 100 MSMEs (100%) have lower scores after macroeconomic factors are included (*negative ranks*), no MSMEs have an increased score (*positive ranks* = 0), and no scores are the same (*ties* = 0). The test statistic is Z = −8.697 with Asymp. Sig. (2-tailed) = 0.000 (Table 3). Since Sig. < 0.05, the null hypothesis is rejected: inflation, deposit interest rates, and the rupiah exchange rate significantly depress the financial health scores of MSMEs. The effect size calculated from Z/√N is r = 0.87, which is considered large.

Table 3. Wilcoxon Signed-Rank Test Results (InEx_ZScore − In_ZScore)

Difference group	N	*Mean rank*	*Sum of ranks*
Negative ranks (InEx < In)	100	50.50	5,050.00
Positive ranks (InEx > In)	0	0.00	0.00
Ties (InEx = In)	0	–	–
Total	100		

Source: SPSS data processing results. Z = −8.697; Asymp. Sig. (2-tailed) = 0.000; based on *positive ranks* .

The significance of these findings lies in their direction and consistency. The internal data-based assessment (Z″-Internal) proved overly optimistic because it failed to account for external pressures, and no MSME is immune to them. This finding aligns with Tinoco and Wilson (2013) that macroeconomic variables carry additional information in assessing financial distress. Two cautions are worth noting. First, the Wilcoxon test only answers whether

differences exist and in what direction, not the magnitude of the decline in scores for each MSME. Second, because all differences are negative, this pattern may be partly mechanical, depending on how macroeconomic factors are incorporated into the scores (e.g., as adjustments that apply equally to all units). Further research should examine the sensitivity of the scores to each macroeconomic variable separately.

Feature selection and machine learning model performance

Feature selection with *Random Forest* reduced 33 features to 11 whose *importance* exceeded the average of 3.03% (Table 4). D1 (11.75%) and D2 (10.03%) occupied the top two positions, with a wide margin over the next feature (inventory, 5.29%). In other words, these two macro/regional indicators carry roughly twice the share of difference as any other financial feature, while the remaining nine features come from the MSME balance sheet and profit and loss.

Table 4. Eleven selected features based on Random Forest *feature importance*

No	Feature	*Importance* (%)	No	Feature	*Importance* (%)
1	D1	11.75	7	Total current assets	4.30
2	D2	10.03	8	Annual net profit	4.23
3	Supply	5.29	9	Cash and cash equivalents	3.49
4	Annual net sales	5.19	10	EBIT	3.45
5	Total assets	5.13	11	Total current liabilities	3.39
6	Total equity/capital	4.59			

Source: Python data processing results (*scikit-learn*). Selection threshold: *importance* > 0.0303.

Table 5 compares the four algorithms on 30 test datasets. Gradient Boosting excels in accuracy (0.80), precision (0.83), *recall* (0.71), and F1-score (0.77), while Random Forest excels in AUC (0.89 versus 0.85). The Gradient Boosting confusion matrix shows 14 *true negatives* , 2 *false positives* , 4 *false negatives* , and 10 *true positives* .

Table 5. Performance of four classification algorithms on test data (n = 30)

Algorithm	Accuracy	Precision	*Recall*	F1-score	AUC
Gradient Boosting	**0,80**	**0,83**	**0,71**	**0,77**	0,85
Random Forest	0,77	0,82	0,64	0,72	**0,89**
Logistic Regression	0,77	0,82	0,64	0,72	0,84
Naive Bayes	0,67	1,00	0,29	0,44	0,78

Source: Python data processing results (*scikit-learn*). Bold letters mark the highest values for each metric.

Synthesizing these figures leads to four points. First, the choice of Gradient Boosting was based on metric balance, not absolute dominance: its advantage over Random Forest was equivalent to one test sample (24 versus 23 correct predictions), so the two models should be considered comparable in performance. Second, for an early warning system, the most costly

error is *a false negative* , that is, a risky MSME that is otherwise healthy. Gradient Boosting missed 4 out of 14 risky MSMEs (28.6%), and its *recall* was significantly better than Naive Bayes, which caught only 29% of risky MSMEs despite a precision of 1.00. Naive Bayes is very conservative: every warning is correct, but most risks go undetected, making it unsuitable as an early warning tool. Third, 80% accuracy on 30 test data sets has substantial uncertainty; the 95% Wilson confidence interval calculated from 24 correct predictions out of 30 ranges from approximately 62.7% to 90.5%. Fourth, this accuracy rate is in the same range as the 82–83% reported by Shetty et al. (2022) on Belgian MSMEs, and the direction of the *boosting method's superiority* over the linear model is in line with the findings of Correa (2023) and Káčer and Alexy (2025), although the data scale of this study is much smaller.

Global interpretation with SHAP

The SHAP ranking in Table 6 places D2 (1.3089), total current assets (0.7944), D1 (0.7776), EBIT (0.7280), and total current liabilities (0.5432) as the five features with the largest average influence on the Gradient Boosting output. The average influence of D2 is approximately 1.6 times that of the second-ranked feature. This ranking differs slightly from the *Random Forest feature selection* , which places D1 above D2. This difference is understandable since the two measures are calculated differently (impurity reduction in Random Forest and Shapley value attribution in Gradient Boosting), but both methods agree that D1 and D2 are the primary determinants.

Table 6. Top five features according to SHAP and their influence direction patterns

Ranking	Feature	*Mean* |SHAP|	Directional pattern of influence (Figure 1)
1	D2	1,3089	High values → negative SHAP (healthy); low values → positive SHAP up to about +2.7
2	Total current assets	0.7944	Mix on both sides; not monotonous
3	D1	0.7776	High value → negative SHAP; three low values → SHAP around +3.3 to +4.0
4	EBIT	0.7280	Low values → negative SHAP; high values → positive SHAP (counterintuitive)
5	Total current liabilities	0.5432	Mostly close to zero; some low values → positive SHAP up to about +3.8

Source: Python data processing results (SHAP *TreeExplainer*) and reading *the summary plot* (Figure 1).

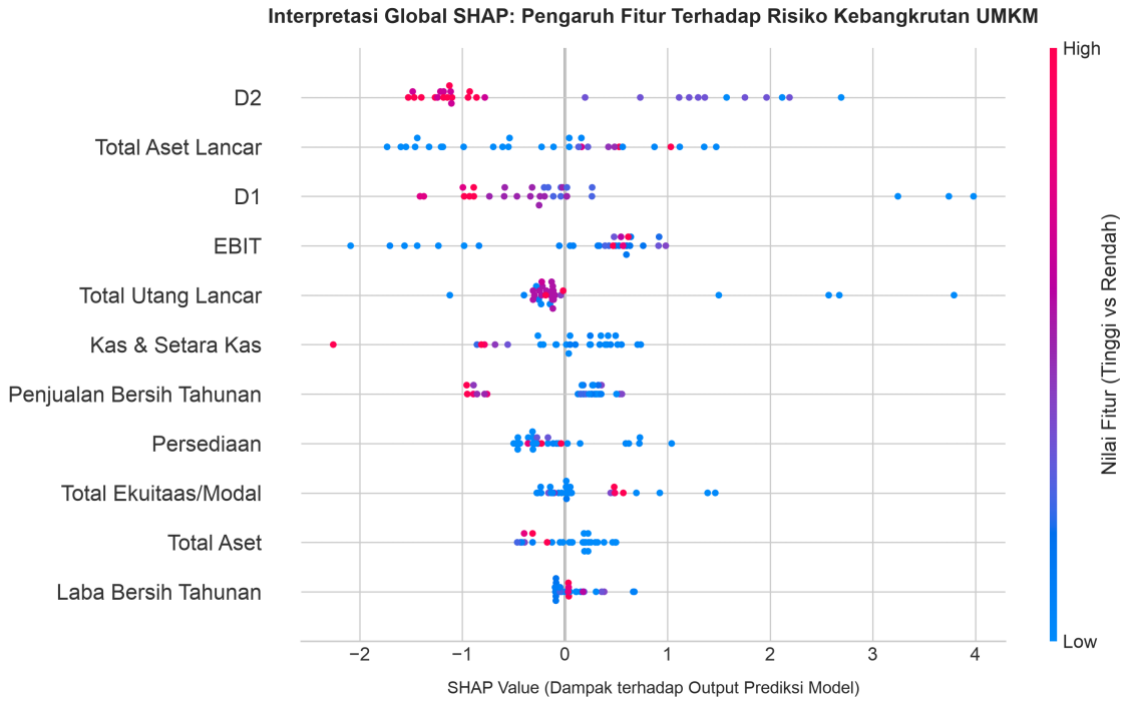


Figure 1. SHAP summary: the influence of eleven features on predicting the risk of bankruptcy of MSMEs

Figure 1 shows the direction of influence of each feature. In D2, high-value points (red) are concentrated at negative SHAP values of around −1.5 to −0.8, while low-value points (blue) are spread out to the positive side, reaching around +2.7. This means that a high D2 value decreases risk, while a low D2 value increases it. D1 has a similar pattern: high values shift toward healthy levels, while three low-value points reach SHAP values of around +3.3 to +4.0. These extreme points are what give D1 its significant influence on some MSMEs, even though its absolute average is below total current assets. Cash and cash equivalents and annual net sales show patterns consistent with economic logic: high values (red) are on the negative side, so high cash acts as a liquidity buffer, and high sales shield MSMEs from risk. In total current assets, the blue and red points are mixed on both sides, so their influence is not monotonic and depends on the combination with other features.

The EBIT pattern is noteworthy because it is counterintuitive. Low EBIT points (blue) are on the negative side of SHAP (up to around −2.1), while high values (red and purple) are on the positive side (around +0.3 to +1.0). In other words, the Gradient Boosting model associates lower EBIT with predicted health. The narrative that adequate EBIT keeps MSMEs healthy is not supported by this graph. Three hypotheses can be tested: (1) the risk label is formed from ten self-reported indicators of business stress, not from financial statements, so larger and more profitable businesses may report more stress indicators; (2) the financial variables are highly correlated (EM correlations of net income with EBIT are 0.997, total assets with equity are 0.995, and inventory with total current assets are 0.911), so SHAP attribution is split among twin features and its direction is unstable; and (3) the small sample size makes the

pattern likely to originate from a few observations. Until that assumption is tested, EBIT should not be interpreted as a cause of health or vulnerability.

Overall, SHAP indicates that the vulnerability of MSMEs in Kupang City is influenced by macro/regional indicators D2 and D1 as well as by the internal liquidity position (total current assets, cash, and current liabilities). This is consistent with the Wilcoxon test finding that external factors depress MSME health, although the two analyses use different outcome measures (Z″ scores and risk labels from indicators E1–E10), so the agreement is thematic and not causal evidence.

Local interpretation with LIME

Table 7 and Figures 2–4 illustrate the predictions for the three MSMEs. MSME #1 is predicted to be healthy with a 98.5% probability (1.5% risk), and its actual status is indeed healthy. The contribution driving toward health amounts to approximately −0.61, primarily from EBIT ≤ −0.62 (−0.24), D2 in the range of −0.60 to 0.26 (−0.20), and D1 (−0.07). The pressure toward risk amounts to approximately +0.29, primarily from low cash and cash equivalents (+0.09) and low equity, sales, profits, and total assets (+0.03 to +0.06, respectively). The net difference yields a confident prediction of health.

Table 7. Summary of local LIME explanations for three MSMEs

MSMEs	Prediction	Total towards risk	Total towards health	Main drivers (LIME weight)
#1	Healthy (98.5%)	+0.29	−0.61	EBIT (−0.24), D2 (−0.20), D1 (−0.07)
#2	Risky	+0.68	−0.42	D2 (+0.35), total current assets (+0.17), EBIT (−0.21)
#3	Healthy	+0.14	−0.75	Total current assets (−0.21), D2 (−0.18), sales (−0.16), D1 (−0.16)

Source: Python (LIME) data processing results. Negative contributions are in the healthy direction and positive contributions are in the risky direction; the number is the sum of the top ten features.

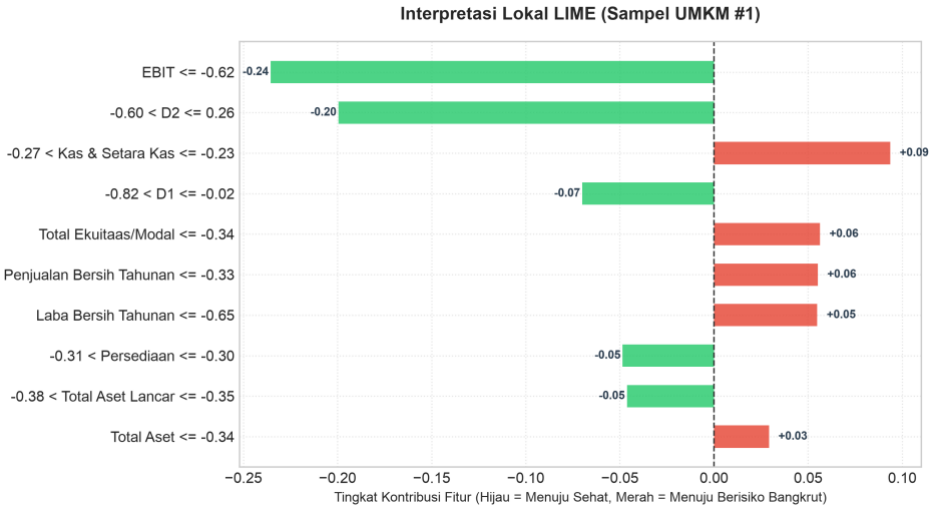


Figure 2. LIME local explanation for MSMEs #1

MSME #2 is predicted to be at risk of bankruptcy. The main causes are D2 ≤ -0.60 which contributes +0.35, followed by total current assets in the range of -0.35 to -0.24 (+0.17), low annual net sales (+0.06), low cash and cash equivalents (+0.05), and low annual net profit (+0.05), totaling +0.68. The balancing factors in the form of EBIT (-0.21), D1 (-0.08), equity (-0.07), inventory (-0.04), and current liabilities (-0.02) only total -0.42. When compared with MSME #1, EBIT is in the same range (≤ −0.62) and D1 is in the same range, so the difference in prediction mainly comes from two features: D2 and total current assets, which together contribute +0.52 of the +0.68 total risk pressure.

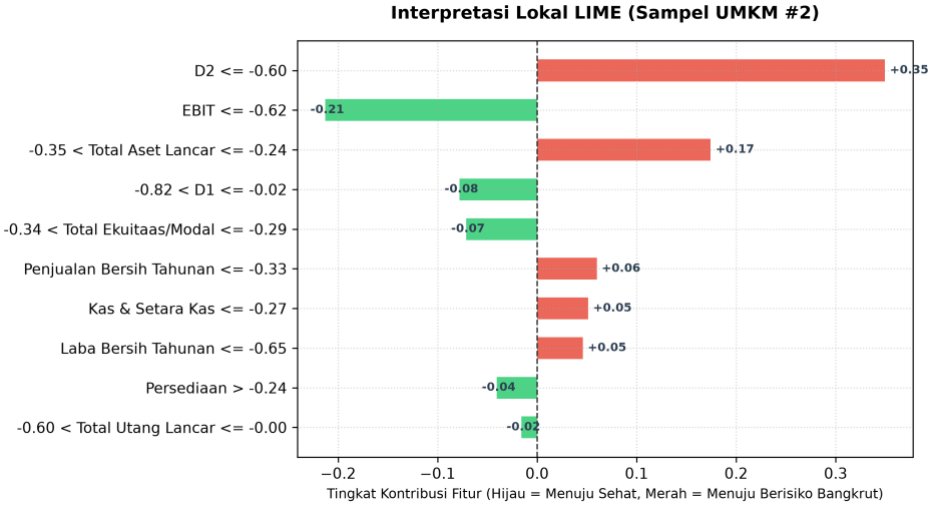


Figure 3. LIME local explanation for MSMEs #2

MSME #3 is predicted to be healthy, with a push towards health of -0.75 compared to risk pressure of only +0.14. The biggest drivers are total current assets ≤ -0.38 (-0.21), D2 in the range of -0.60 to 0.26 (-0.18), annual net sales above -0.05 (-0.16), and D1 in the range of -0.02 to 0.77 (-0.16). Small pressures come from cash and cash equivalents (+0.06) and equity (+0.06).

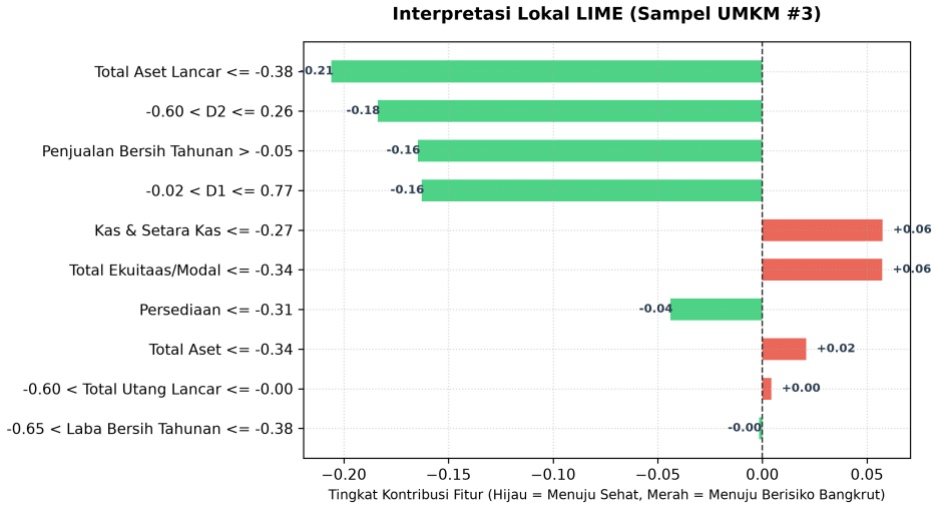


Figure 4. LIME local explanation for MSMEs #3

These three explanations reveal something that is not apparent in the global rankings. First, D2 acts like a threshold: in the range of −0.60 to 0.26, it pushes toward health (SMEs #1 and #3), while below −0.60, it becomes the largest driver of risk (SME #2). A D2 value less than approximately 0.6 standard deviations below the mean thus appears to be a critical point for the model. Second, the same feature can act differently across its range; total liquid assets push toward health in SME #3 (−0.38 or lower) but toward risk in SME #2 (−0.35 to −0.24), which is consistent with the non-monotonicity of the SHAP distribution. Third, low cash and equity appear as risk pressures in SMEs that are still predicted to be healthy, signaling latent vulnerability that is not apparent when only the final label is read. The LIME explanation is local and based on a linear surrogate model, so its weights can change with different perturbation samples; he explains how the model reasons, not the cause and effect of actual work.

Synthesis and implications

The three layers of analysis provide a complementary picture. The Z″ score indicates the polarization of financial health and the fragility of ratios in businesses with almost no debt. The Wilcoxon test indicates that macroeconomic factors shift the assessment downward for all MSMEs. *Machine learning models* using SHAP and LIME show that indicators D1 and D2, along with short-term liquidity, are the main determinants of risk prediction, and that their influence depends on the threshold value. The implications for stakeholders are practical. For local governments and fostering institutions, monitoring indicators D1 and D2 and their critical thresholds can serve as a basis for determining mentoring priorities. For banks and financing institutions, LIME output provides a verifiable rationale for each decision, not just a score. For business owners, cash management, current assets, and sales are internal factors that can be addressed independently.

CONCLUSION AND SUGGESTIONS

This study addresses all four objectives. Missing data patterns across 100 MSMEs in Kupang City do not exhibit any systematic biases in the analysis (Sig. 0.587). Macroeconomic factors significantly reduced the financial health scores of all MSMEs (Z = −8.697; Sig. 0.000), making an assessment based solely on internal data inadequate. Of the four algorithms, Gradient Boosting provided the best balance of metrics (80% accuracy, 0.71 *recall* , 0.85 AUC) and, together with XAI, identified risk determinants: D2, total current assets, and D1 dominated globally, while at the business unit level, the combination of low D2 and limited current assets differentiated risky MSMEs from healthy ones. Thus, a transparent predictive model is suitable as the basis for a regional MSME early warning system.

These results need to be interpreted within the limitations of the study. (1) The sample size was only 100 MSMEs from one city and the test data was only 30 MSMEs with a single data split, so the model performance cannot be generalized and small differences between algorithms cannot be considered definitive. (2) Mean imputation, standardization, and feature selection were performed on all data before the train-test split and the algorithm used default parameters without cross-validation, so the performance is potentially overly optimistic. (3) The risk labels were derived from the business stress indicators E1–E10 with a threshold of four "Yes" answers, not from actual bankruptcy events, and the threshold is conventional. (4) The EM algorithm did not converge in 25 iterations and the Z″ score contained extreme values due to a very small denominator. (5) High correlations between financial variables, macro adjustments that have not been described per variable, and the meaning of D1 and D2 that have not been substantively described limit the interpretation of SHAP and LIME, including the counterintuitive direction of the EBIT effect.

Based on this, further research is recommended to: enlarge the sample and regional coverage; use cross-validation, hyperparameter tuning, and apply preprocessing only to training data; compare with other boosting algorithms; use longitudinal data with real bankruptcy events as targets; handle extreme ratios with *winsorizing* or alternative ratios; and test the stability of attribution on correlated features. For practitioners, local governments and financial institutions are advised to build early warning monitoring boards based on the key features and D2 thresholds found, equipped with financial recording assistance to ensure that accounts receivable and payable are no longer empty. MSMEs are advised to maintain cash reserves and current assets and monitor sales as independent mitigation measures.

BIBLIOGRAPHY

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609. https://doi.org/10.2307/2325319

Altman, E. I. (2005). An emerging market credit scoring system for corporate bonds. *Emerging Markets Review, 6*(4), 311–323.

Altman, E. I. (2018). Applications of distress prediction models: What have we learned after 50 years from the Z-score models? *International Journal of Financial Studies, 6*(3), 70. https://doi.org/10.3390/ijfs6030070

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management & Accounting, 28*(2), 131–171. https://doi.org/10.1111/jifm.12053

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

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58* , 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Breiman, L. (2001). Random forests. *Machine Learning, 45* (1), 5–32. https://doi.org/10.1023/A:1010933404324

Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics, 57* (1), 203–216. https://doi.org/10.1007/s10614-020-10042-0

Correa, A. (2023). Predicting business bankruptcy in Colombian SMEs: A machine learning approach. *Journal of International Commerce, Economics and Policy, 14*(3). https://doi.org/10.1142/S1793993323500278

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gramegna, A., & Giudici, P. (2021). SHAP and LIME: An evaluation of discriminative power in credit risk. *Frontiers in Artificial Intelligence, 4*, 752558. https://doi.org/10.3389/frai.2021.752558

Káčer, M., & Alexy, M. (2025). SMEs' failure prediction: Are machine learning models superior? *Studies in Economics and Finance, 42*(5), 908–944. https://doi.org/10.1108/SEF-09-2024-0629

Kou, G., Xu, Y., Peng, Y., Shen, F., Chen, Y., Chang, K., & Kou, S. (2021). Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection. *Decision Support Systems, 140*, 113429. https://doi.org/10.1016/j.dss.2020.113429

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global

understanding with explainable AI for trees. *Nature Machine Intelligence, 2*(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

Nguyen, H. H., Viviani, J.-L., & Ben Jabeur, S. (2025). Bankruptcy prediction using machine learning and Shapley additive explanations. *Review of Quantitative Finance and Accounting, 65*(1), 107–148. https://doi.org/10.1007/s11156-023-01192-x

Park, M. S., Son, H., Hyun, C., & Hwang, H. J. (2021). Explainability of machine-learning models for bankruptcy prediction. *IEEE Access, 9*, 124887–124899. https://doi.org/10.1109/ACCESS.2021.3110270

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Romaniyah, & Alfian, N. (2024). Analysis of financial distress prediction using the Altman Z-Score model in MSMEs in Pamekasan. *KRISNA: Accounting Research Collection, 15* (2), 309–316. https://doi.org/10.22225/kr.15.2.2024.309-316

Shetty, S., Musa, M., & Brédart, X. (2022). Bankruptcy prediction using machine learning techniques. *Journal of Risk and Financial Management, 15* (1), 35. https://doi.org/10.3390/jrfm15010035

Tambunan, TTH (2019). Recent evidence of the development of micro, small and medium enterprises in Indonesia. *Journal of Global Entrepreneurship Research, 9* , 18. https://doi.org/10.1186/s40497-018-0140-4

Tinoco, M. H., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International Review of Financial Analysis, 30*, 394–419. https://doi.org/10.1016/j.irfa.2013.02.013

Wahlstrøm, R. R., Becker, L.-K., & Fornes, T. N. (2024). Enhancing credit risk assessments of SMEs with non-financial information. *Cogent Economics & Finance, 12*(1), 2418910. https://doi.org/10.1080/23322039.2024.2418910

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83. https://doi.org/10.2307/3001968

Yeo, W.J., et al. (2025). A comprehensive review on financial explainable AI. *Artificial Intelligence Review, 58* (6), 189. https://doi.org/10.1007/s10462-024-11077-7