JSiI | Jurnal Sistem Informasi
Vol. 13 No. 1 | March 2026, P. 106-114

p-ISSN: 2406-7768
e-ISSN: 2581-2181

ANALYSIS OF THE ACCURACY LEVEL OF FINANCIAL DISTRESS PREDICTION MODELS USING THE NAÏVE BAYES METHOD

Ridho Dwi Maulida[1], Arief Wibowo[2], Selamet Riyadi[3]

[1] *Magister Manajemen, Fakultas Ekonomi dan Bisnis, Universitas Budi Luhur*
[2, 3] *Doktor Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Budi Luhur*

Jl. Ciledug Raya – Kota Jakarta Selatan
[1]dwimaulida456@gmail.com
[2]arief.wibowo@budiluhur.ac.id
[3]selamet.riyadi@budiluhur.ac.id

Abstract

The ability to accurately predict financial distress is crucial for State-Owned Enterprises (SOEs), given their strategic role in maintaining national economic stability. However, existing studies predominantly examine financial distress models in isolation and rely mainly on financial ratios, with limited attention to comparative evaluation under a unified machine learning framework and alternative input structures. This gap limits the understanding of how model performance may vary across different data representations. This study aims to evaluate and compare the predictive performance of four financial distress models Altman Z-Score, Springate S-Score, Zmijewski X-Score, and Grover G-Score by integrating them within a Naïve Bayes classification approach. Using a dataset of 20 Indonesian SOEs listed on the Indonesia Stock Exchange over the 2020–2023 period, this study applies a quantitative comparative method with two types of input variables, namely financial ratios and financial statement account balances. The results show that the Springate S-Score model demonstrates the highest predictive accuracy, achieving 95% when using financial ratios and 82.5% when using account balances. Overall, models based on financial ratios outperform those utilizing raw financial statement data, indicating that structured financial indicators provide more effective signals for classification. The main contribution of this study lies in providing a comprehensive and consistent comparison of multiple financial distress prediction models within a single probabilistic machine learning framework, while also highlighting the impact of different input variable structures on model performance. This study extends the financial distress literature by bridging traditional financial analysis and data mining approaches, and offers practical implications for developing more reliable early warning systems for financial distress in SOEs.

Keywords : *Financial Distress Prediction, Naïve Bayes, Machine Learning*

I. INTRODUCTION

The ability to predict financial distress has become increasingly important, particularly for State-Owned Enterprises (SOEs), given their strategic role in maintaining economic stability. Financial distress not only affects corporate sustainability but may also have broader systemic implications for national economies. Therefore, the development of accurate and reliable prediction models is essential for early warning systems and decision-making processes.

With the advancement of technology, discovering patterns, anomalies, predictions and relationships in data, particularly large scale data has become more manageable [1]. In recent years, the application of machine learning techniques in financial distress prediction has gained significant attention. Various classification methods, such as Decision Trees, Random Forest, Support Vector Machines (SVM) and Naïve Bayes, have been widely applied to improve predictive performance [2]. Among these various data mining techniques, classification is one of the most frequently used methods. Classification is a technique that predicts class labels based on input data [3]. This technique is particularly important and widely used in analyzing and predicting corporate financial performance, as many data models used in this context tend to exhibit relational characteristics [4].

Studies within the last decade indicate that machine learning approaches often outperform traditional statistical models due to their ability to capture complex patterns in financial data. However, the performance of these methods is highly dependent on the quality and structure of input variables.

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

Despite the growing use of machine learning, many prior studies tend to focus on a single financial distress model or a limited set of predictors. Comparative analyses that systematically evaluate multiple established financial distress models such as Altman Z-Score, Springate S-Score, Zmijewski X-Score, and Grover G-Score using a consistent machine learning framework remain relatively limited, particularly in the context of Indonesian SOEs. Furthermore, existing studies often emphasize financial ratios without examining the predictive capability of raw financial statement account balances as alternative inputs.

By focusing on Indonesian SOEs, this study also contributes to the literature by providing empirical evidence in a specific economic context characterized by large-scale enterprises with complex financial structures. The findings are expected to offer both theoretical contributions to financial analytics and practical implications for stakeholders in developing more effective financial distress early warning systems.

This study employs the Naïve Bayes method as the classification technique to analyze corporate financial data. Although several other classification methods exist, such as Decision Trees, Random Forest and Support Vector Machines (SVM), previous research [2] has shown that Naïve Bayes outperforms models like Decision Trees and SVM in terms of accuracy and precision. For this reason, Naïve Bayes is selected for this study due to its ability to provide more accurate and efficient predictions an essential factor in predicting financial distress [5]. The goal is to develop a predictive model that accurately identifies financial distress in line with the study's objectives.

Data mining in general is applied to analyze a company's financial statements. Financial statements are often used in data mining because they reflect key financial indicators and other important information related to a company's financial behavior [6]. Common financial data used to evaluate a company include accounts receivable, net income, assets, and capital structure. The application of data mining techniques to financial statement analysis serves, among other purposes, to assess a company's financial health [7]. Financial health is a critical indicator often used by investors when making investment decisions [7].

Several models have been developed to assess a company's financial health, one of which involves calculating the risk of financial distress. The concept of financial distress was first introduced by Edward I. Altman in 1968. Initially, Altman referred to financial difficulties as a condition in which a company struggles financially, characterized by its inability to pay debts and the reporting of negative net income (loss) [8]. In his research, Altman developed a model to predict financial distress levels in companies, which later became widely known as the Altman Z-Score. Over time, subsequent researchers have continued to develop financial distress prediction models [9].

Due to the continuous evolution of these predictive models, many researchers have investigated financial distress prediction, often exploring the contributing factors and their direct implications. Previous research [10] found that a company experiencing financial distress typically sees a direct decline in its stock price. Another study [11] highlighted that financially distressed firms tend to face solvency issues, where the total assets are insufficient to cover liabilities a situation further exacerbated by a decline in revenue.

In addition, previous studies tend to rely on a single type of input variable, predominantly financial ratios, without exploring whether alternative data representations such as raw financial statement account balances could yield different predictive performance. This indicates a methodological gap, as the choice of input variables may significantly influence classification outcomes, yet remains underexplored in the literature.

Moreover, inconsistencies in findings across studies regarding the relative accuracy of financial distress models suggest that no single model can be universally considered superior. These variations may be attributed to differences in sample characteristics, industry context and analytical approaches, highlighting the need for a more standardized and comparative evaluation.

Some studies have attempted to predict financial distress using only specific proxies such as the Springate S-Score and Zmijewski X-Score one such study being conducted by [12], [13], [14], [15]. Based on these observations, it can be concluded that studies comprehensively comparing the accuracy levels of all major financial distress prediction models remain limited within the academic literature.

Therefore, this study addresses these gaps by conducting a comprehensive comparison of four widely used financial distress prediction models Altman Z-Score, Springate S-Score, Zmijewski X-Score and Grover G-Score within a unified classification framework using the Naïve Bayes algorithm. Unlike prior research, this study not only compares multiple models simultaneously but also evaluates the impact of different input structures, namely financial ratios and financial statement account balances, on predictive performance.

The importance of this research lies in its potential to provide a more robust and holistic understanding of financial distress prediction, both methodologically and practically. From a methodological perspective, it contributes to the literature by integrating traditional financial models with a machine learning approach under consistent evaluation metrics. From a practical standpoint, the findings can assist investors, analysts and policymakers in selecting the most appropriate prediction model based on data availability and desired accuracy.

Thus, this study seeks to answer the following research question: which of the four financial distress prediction models demonstrates the highest level of predictive accuracy when analyzed using the Naïve Bayes classification method?

II. RESEARCH METHOD

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

This study employs a quantitative approach using a knowledge discovery in database (KDD) process to predict financial distress in Indonesian State-Owned Enterprises (SOEs). The classification technique used in this research is the Naïve Bayes algorithm due to its simplicity, efficiency and strong performance in handling relatively small datasets.

The dataset consists of financial statements from 20 SOEs listed on the Indonesia Stock Exchange over the period 2020–2023, resulting in a total of 80 observations. The data includes both financial ratios and raw financial statement account balances, which are used as input variables in the prediction models.

Data preprocessing is conducted prior to model implementation. This stage includes data cleaning to handle missing or inconsistent values, data transformation to ensure uniform variable formats and normalization where necessary to improve model performance [14]. Additionally, the financial variables are calculated based on the formulas of the Altman Z-Score, Springate S-Score, Zmijewski X-Score, and Grover G-Score models. The formulas applied for each financial distress prediction model in this study are as follows:

a) Altman Z-Score

$$Z - Score = 6.56\,x\frac{Working\ Capital}{Total\ Asset} + 3.267\,x\ \frac{Retained\ Earnings}{Total\ Asset} + 6.72\,x\ \frac{EBIT}{Total\ Asset} + \ 1.05\ \frac{MVE}{Total\ Liabilities}$$

Source : [8]

b) Springate S-Score

$$S - Score = 1.03\,x\frac{Working\ Capital}{Total\ Asset} + 3.07\,x\ \frac{EBIT}{Total\ Asset} + 0.66\,x\ \frac{EBT}{Current\ Liabilities} + 0.4\,x\ \frac{Sales}{Total\ Asset}$$

Source : [16]

c) Zmijewski X-Score

$$X - Score = -4.3 - 4.5\,x\frac{Net\ Income}{Total\ Asset} + 5.7\,x\ \frac{Total\ Liabilities}{Total\ Asset} - 0.004\,x\ \frac{Current\ Assets}{Current\ Liabilities}$$

Source: [17]

d) Grover G-Score

$$G - Score = \ 1.650\,x\frac{Working\ Capital}{Total\ Asset} + 3.404\,x\ \frac{EBIT}{Total\ Asset} - 0.016\,x\ \frac{Net\ Income}{Total\ Asset} + 0.057$$

Source : [18]

The research model employed in this study is illustrated in the following diagram:

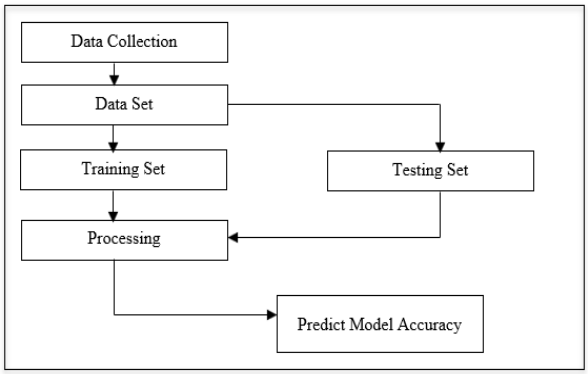


Figure 1. Analysis Steps.

Based on the illustrated framework, the data analysis process in this study consists of several systematic stages to ensure transparency and replicability. The initial stage involves collecting secondary data in the form of annual financial statements from state-owned enterprises listed on the Indonesia Stock Exchange for the period 2020–2023. The collected data are then compiled into a structured dataset for further processing.

The next stage is data preprocessing, which is conducted to improve data quality and model performance. This process includes (1) data cleaning, where incomplete, inconsistent, or missing values are identified and handled using data exclusion or imputation techniques where appropriate; (2) data transformation, where raw financial statement accounts are standardized into a consistent format; and (3) feature construction, where financial ratios are calculated based on the formulas of the Altman Z-Score, Springate S-Score, Zmijewski X-Score, and Grover G-Score models. For the ratio-based approach, variables are expressed in proportional form, while for the financial statement account approach, raw account values are retained as model inputs.

Following preprocessing, the dataset is partitioned into training and testing sets using a 70:30 split ratio. Specifically, 70% of the observations are used to train the model, while the remaining 30% are reserved for out-of-sample testing to evaluate model generalization performance. To further enhance the robustness and reduce potential sampling bias, a 10-fold cross-validation technique is applied within the training process. In this approach, the training data are divided into ten subsets, where nine subsets are used for model training and one subset for validation, repeated iteratively until all subsets have been used for validation.

The predictive modeling stage employs the Naïve Bayes classification algorithm, which is selected due to its efficiency and suitability for probabilistic classification problems [19]. Model training and evaluation are implemented using RapidMiner software, utilizing its built-in operators for data preprocessing, model training, and validation.

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

Model performance is evaluated using multiple metrics to provide a comprehensive assessment. These include accuracy to measure overall correctness, precision to evaluate the reliability of positive classifications and recall to assess the model's sensitivity in detecting actual classes [4]. Finally, the results obtained from each model are analyzed and compared to determine the most effective financial distress prediction model [3].

III. RESULT AND DISCUSSION

Result

Based on the data collection process, out of the 27 state-owned enterprises (SOEs) listed on the Indonesia Stock Exchange, only 20 companies meet the sample criteria and will be used in this study. The data used in this research is sourced from the financial statements of these companies for the period 2020-2023, which were obtained through the official website of the Indonesia Stock Exchange. The results of the accuracy calculations for each model are all computed using RapidMiner software will be presented in the following figures:

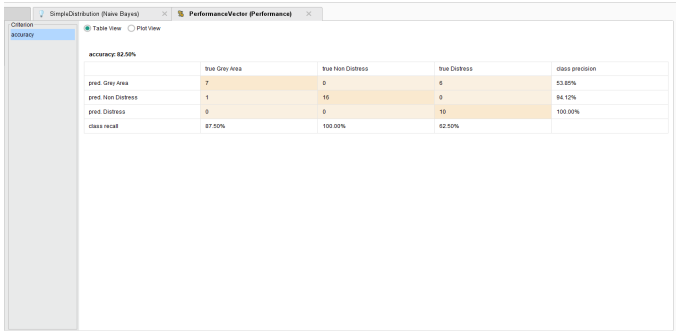


Figure 2. Accuracy Prediction Testing of Ratio-Based Models in the Altman Z-Score Formula Using Naïve Bayes Classifier

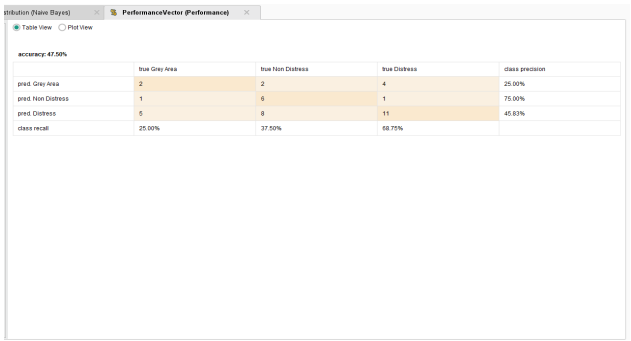


Figure 3. Accuracy Prediction Testing of the Model Based on Financial Statement Account Balances Used to Calculate the Altman Z-Score with Naïve Bayes Classifier

Table 1. Altman Z-Score Result

Altman Z-Score	Ratio-Based	Financial Statement Account
Accuracy Percentage		
	82.5%	47.5%
Class Precission		
Grey Area	53.85%	25%
Non Distress	94.12%	75%
Distress	100%	45.83%
Class Recall		
Grey Area	87.5%	25%
Non Distress	100%	37.5%
Distress	62.5%	68.75%

The analysis of accuracy models applies two different calculation of the Altman Z-Score framework, the Financial Statement Account and the Ratio-Based model.

The overall accuracy of the Ratio-Based model (82.5%) is considerably higher than that of the Financial Statement Account approach (47.5%). This finding suggests that the transformation of financial data into ratios enhances the model's ability to capture the relative financial condition of firms. In contrast, the use of raw financial statement accounts may limit the model's effectiveness, as it does not adequately reflect proportional relationships among financial variables.

In terms of class precision, the ratio-based model demonstrates consistently superior performance across all categories. The precision for the grey area classification is 53.85%, compared to 25% for the financial statement account approach, indicating improved capability in identifying firms within uncertain conditions. For the non-distress category, the ratio-based model achieves a precision of 94.12%, whereas the financial statement account approach records 75%, suggesting a higher level of reliability in classifying financially healthy firms. In the distress category, the ratio-based model attains perfect precision (100%), implying the absence of false positive classifications, while the financial statement account approach only achieves 45.83%, reflecting a relatively high misclassification rate.

From the perspective of class recall, the ratio-based model also outperforms the financial statement account approach in most categories. The recall for the grey area category reaches 87.5%, significantly higher than the 25% observed in the financial statement account approach, indicating stronger sensitivity in detecting firms within the grey zone. Similarly, in the non-distress category, the ratio-based model achieves a recall of 100%, compared to 37.5% for the financial statement account approach. However, in the distress category, the financial statement account approach is 68.75% which is slightly exceeds the ratio-based model (62.5%), suggesting that on distress category, financial statement account is more sensitive in capturing a greater proportion of actual distressed firms.

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

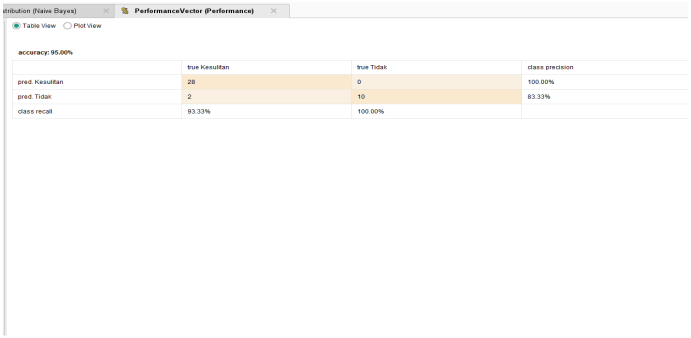


Figure 4. Accuracy Prediction Testing of Ratio-Based Models in the Springate S-Score Formula Using Naïve Bayes Classifier

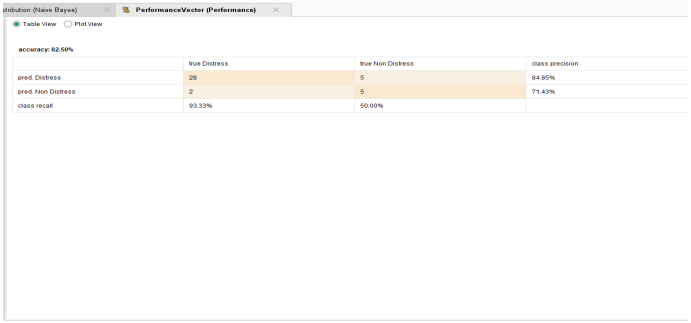


Figure 5. Accuracy Prediction Testing of the Model Based on Financial Statement Account Balances Used to Calculate the Springate S-Score with Naïve Bayes Classifier

Table 2. Springate S-Score Result

Springate S-Score	Ratio-Based	Financial Statement Account
Accuracy Percentage		
	95%	82.5%
Class Precission		
Non Distress	83.33%	71.43%
Distress	100%	84.85%
Class Recall		
Non Distress	100%	50%
Distress	93.33%	93.33%

The analysis of accuracy models applies two different calculation approaches within the Springate S-Score framework, namely the financial statement account approach and the ratio-based model.

The overall accuracy of the ratio-based model (95%) is higher than that of the financial statement account approach (82.5%). This finding indicates that transforming financial data into ratio form improves the model's ability to represent the relative financial condition of firms. In contrast, the use of raw financial statement accounts may reduce the effectiveness of the model, as it does not fully capture proportional relationships among financial variables.

In terms of class precision, the ratio-based model demonstrates better performance across all categories. For the non-distress category, the ratio-based model achieves a precision of 83.33%, compared to 71.43% for the financial statement account approach, indicating greater reliability in identifying financially healthy firms. In the distress category, the ratio-based model attains perfect precision (100%), implying that all firms classified as distressed are correctly identified, whereas the financial statement account approach achieves 84.85%, reflecting a higher likelihood of misclassification.

From the perspective of class recall, the ratio-based model also shows stronger performance in the non-distress category, with a recall of 100% compared to 50% for the financial statement account approach. This suggests that the ratio-based model is more effective in capturing all actual non-distressed firms. In the distress category, both approaches demonstrate equal performance, with a recall value of 93.33%, indicating that both models have a similar ability to identify the majority of distressed firms.

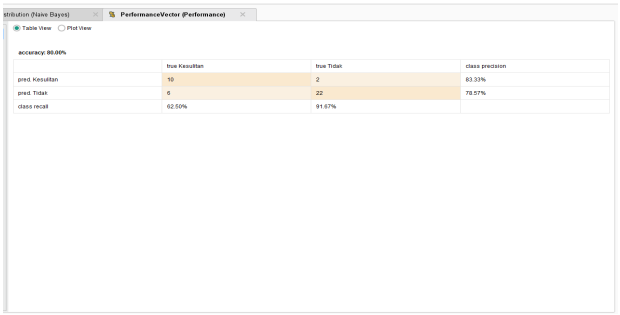


Figure 6. Accuracy Prediction Testing of Ratio-Based Models in the Zmijewski X-Score Formula Using Naïve Bayes Classifier

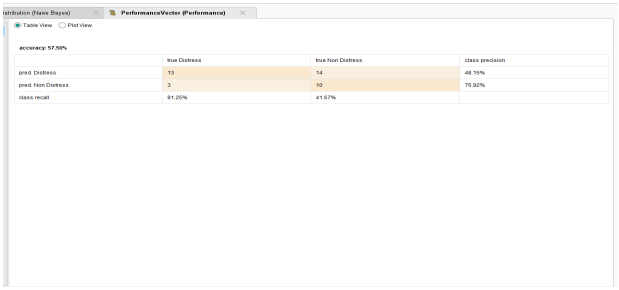


Figure 7. Accuracy Prediction Testing of the Model Based on Financial Statement Account Balances Used to Calculate the Zmijewski X-Score with Naïve Bayes Classifier

Table 3. Zmijewski X-Score Result

Zmijewski X-Score	Ratio-Based	Financial Statement Account
Accuracy Percentage		
	80%	57.5%
Class Precission		
Non Distress	78.57%	76.92%

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

Distress	83.33%	48.15%
	Class Recall	
Non Distress	91.67%	41.67%
Distress	52.5%	81.25%

The analysis of accuracy models applies two different calculation approaches within the Zmijewski X-Score framework, namely the financial statement account approach and the ratio-based model.

The overall accuracy of the ratio-based model (80%) is higher than that of the financial statement account approach (57.5%). This indicates that transforming financial statement data into ratio form enhances the model's ability to capture the relative financial condition of firms. In contrast, the use of raw financial statement accounts appears less effective in representing proportional relationships among financial variables, leading to lower classification accuracy.

In terms of class precision, the ratio-based model shows generally better performance, particularly in the distress category. For the non-distress category, the ratio-based model achieves a precision of 78.57%, which is slightly higher than the financial statement account approach (76.92%), indicating relatively similar capability in identifying financially healthy firms. However, in the distress category, the ratio-based model records a significantly higher precision (83.33%) compared to 48.15% for the financial statement account approach, suggesting that the ratio-based model is more reliable in correctly classifying distressed firms and minimizing false positives.

From the perspective of class recall, different patterns emerge across categories. In the non-distress category, the ratio-based model achieves a recall of 91.67%, substantially higher than the 41.67% recorded by the financial statement account approach, indicating stronger ability to capture all actual non-distressed firms. Conversely, in the distress category, the financial statement account approach (81.25%) outperforms the ratio-based model (52.5%), suggesting that it is more sensitive in identifying a greater proportion of actual distressed firms.

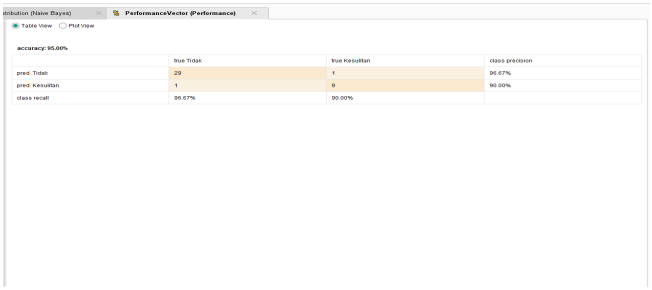


Figure 8. Accuracy Prediction Testing of Ratio-Based Models in the Grover G-Score Formula Using Naïve Bayes Classifier

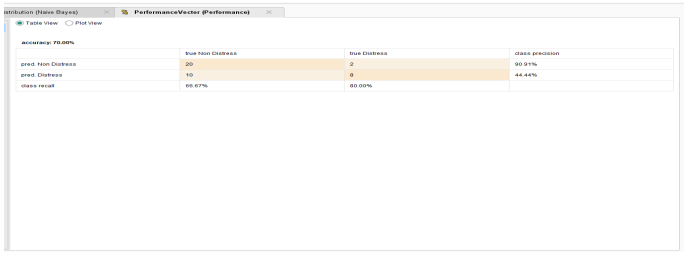


Figure 9. Accuracy Prediction Testing of the Model Based on Financial Statement Account Balances Used to Calculate the Grover G-Score with Naïve Bayes Classifier

Table 4. Grover G-Score Result

Grover G-Score	Ratio-Based	Financial Statement Account
	Accuracy Percentage	
	95%	70%
	Class Precission	
Non Distress	96.67%	90.91%
Distress	90%	44.44%
	Class Recall	
Non Distress	96.67%	66.67%
Distress	90%	80%

The analysis of accuracy models applies two different calculation approaches within the Grover G-Score framework, namely the financial statement account approach and the ratio-based model.

The overall accuracy of the ratio-based model (95%) is substantially higher than that of the financial statement account approach (70%). This finding indicates that transforming financial data into ratio form enhances the model's ability to more accurately represent the financial condition of firms. In contrast, the use of raw financial statement accounts may limit the model's effectiveness due to the lack of proportional standardization among variables.

In terms of class precision, the ratio-based model demonstrates superior performance across both categories. For the non-distress category, the ratio-based model achieves a precision of 96.67%, compared to 90.91% for the financial statement account approach, indicating a higher level of reliability in identifying financially healthy firms. In the distress category, the difference is more pronounced, where the ratio-based model records a precision of 90%, while the financial statement account approach only achieves 44.44%, suggesting a significantly higher misclassification rate in the latter.

From the perspective of class recall, the ratio-based model also outperforms the financial statement account approach in both categories. In the non-distress category, the ratio-based model achieves a recall of 96.67%, compared to 66.67% for the financial statement account approach, indicating stronger capability in capturing all actual non-distressed firms. Similarly, in the distress category, the ratio-based model records a recall of 90%, which is higher than the 80% achieved

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

by the financial statement account approach, suggesting better sensitivity in identifying distressed firms.

Table 5. Summary of Financial Distress Prediction Accuracy Using Naïve Bayes Classifier

No.	Calculation Model	Accuracy Based on Financial Ratios (%)	Accuracy Based on Financial Statement Account Balances (%)
1	Altman Z-Score	82.5	47.5
2	Springate S-Score	95.0	82.5
3	Zmijewski X-Score	80.0	57.5
4	Grover G-Score	95.0	70.0

Table 5 presents a comparative summary of the prediction accuracy of four financial distress models using the Naïve Bayes classifier. The results indicate that ratio-based models consistently outperform models based on financial statement account balances across all prediction methods. Among the four models, the Springate S-Score and Grover G-Score demonstrate the highest accuracy when using financial ratios, each achieving an accuracy rate of 95 percent. However, when account balances are used as input variables, the Springate S-Score maintains relatively high predictive performance at 82.5 percent, indicating greater robustness compared to the other models.

Discussion

The findings of this study provide a more comprehensive evaluation of financial distress prediction models by incorporating additional performance metrics beyond accuracy using class precision and recall. These metrics class precision and recall offer deeper insights into the classification behavior of each model, particularly in distinguishing between non-distress and distress conditions.

The empirical results indicate that the representation of financial data plays a critical role in determining model performance. Models based on financial ratios consistently outperform those using raw financial statement account balances across all evaluation metrics. This can be explained by the ability of financial ratios to capture proportional relationships between key financial variables, such as liquidity, profitability and leverage. Ratios provide a normalized and standardized form of data, which enhances comparability across firms and improves the learning process of the Naïve Bayes classifier. In contrast, raw financial statement figures are expressed in absolute terms, which may introduce scale bias and reduce the model's ability to generalize across different firms.

A deeper evaluation using precision reveals that ratio-based models are more effective in minimizing false positive classifications, particularly in the distress category. This indicates that when a firm is classified as distressed, there is a higher probability that the classification is correct. Such reliability is crucial for stakeholders, including investors and creditors, who depend on accurate signals to make financial decisions. On the other hand, recall analysis shows that while ratio-based models generally perform better in identifying non-distressed firms, models based on financial statement account balances tend to exhibit relatively higher sensitivity in certain distress classifications. This implies that the account-based approach is more effective in capturing a larger proportion of actual distressed firms, although this increased sensitivity is accompanied by a greater likelihood of misclassification.

Based on the summary of testing results described on Table 5, this research indicate that the Springate S-Score and Grover G-Score model provides the most precise models to detect financial distress conditions within companies with an accuracy rate of 95 percent. While utilizing financial ratios in the model calculation, Springate S Score provides the most precise models achieving 82.5 percent to predict financial distress.

The high accuracy of the Springate S-Score and Grover G-Score models can be attributed to their emphasis on key financial ratios that capture both profitability and operational efficiency. These models incorporate variables such as earnings relative to total assets and asset turnover, which reflect how effectively firms utilize their resources to generate income [20]. In addition, their parsimonious structure, relying on a limited number of informative variables, reduces model complexity and mitigates the risk of overfitting within the Naïve Bayes framework [21]. As Naïve Bayes assumes conditional independence among predictors, the use of less correlated variables enables more stable probabilistic estimation. In contrast, the inclusion of numerous financial indicators or raw account values increases the likelihood of redundancy and multicollinearity, which may weaken classification performance.

In contrast, the Altman Z-Score and Zmijewski X-Score models rely more heavily on broader financial structures, including leverage and liquidity components. While these variables are important, they may not fully capture the operational performance of firms, especially in dynamic business environments [21]. As a result, these models tend to be more sensitive to financial structure rather than efficiency, which can reduce their discriminative power when compared to models that incorporate more performance-oriented ratios [13].

From a methodological standpoint, the inclusion of precision and recall as evaluation metrics strengthens the validity of the results by providing a more nuanced understanding of model effectiveness. These metrics highlight the trade-off between accuracy, reliability and sensitivity in financial distress prediction [22]. While ratio-based models offer higher overall accuracy and precision, financial statement account-based models may still contribute value in contexts where identifying as many potential distress cases as possible is prioritized.

This study contributes to the financial distress prediction literature by showing that models based on financial ratios tend to achieve better predictive performance than those using raw financial statement data. This finding supports the view that ratio-based measures, as normalized indicators, are more effective in capturing firm conditions and improving

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

comparability across observations. In addition, the results indicate that models emphasizing efficiency and profitability ratios, such as the Springate S-Score and Grover G-Score, are better at distinguishing between distressed and non-distressed firms [23], [24].

This finding is supported by research conducted by [20], which asserts that the Springate S-Score model has consistently demonstrated the highest accuracy compared to other models. The advantage of the asset turnover ratio lies in its capacity to offer comprehensive insights into a company's asset productivity. It indicates how effective a company employs both fixed and current assets to generate sales. Since this ratio reflects operational efficiency, it can aid in the early detection of potential financial distress. Companies with low asset efficiency may face cash flow difficulties due to insufficient revenue generation to cover operational costs or other obligations. Conversely, a high ratio indicates effective asset management and a stronger capacity to maintain financial stability.

IV. CONCLUSION

Based on the predictive evaluation employing the Naïve Bayes classification method, this study confirms that the Springate S-Score and Grover G-Score models provide the highest predictive performance among the examined models in detecting corporate financial distress. Both models achieve an accuracy rate of 95 percent when financial ratios are used, while the Springate S-Score remains the most robust model under the financial statement account approach, with an accuracy of 82.5 percent. These findings indicate that model performance is not only influenced by the choice of algorithm, but also by the structure of the model and the representation of financial data [25].

The results further demonstrate that the use of financial ratios significantly enhances predictive capability compared to raw financial statement account balances. This highlights the importance of data transformation in financial distress prediction, as ratios are able to capture proportional relationships and improve comparability across firms. In addition, the inclusion of operational efficiency indicators, particularly asset turnover, contributes to the superior performance of the Springate and Grover models, suggesting that efficiency-based variables provide more informative signals in identifying early signs of financial distress.

From a practical perspective, the findings offer important implications for stakeholders such as investors, creditors, and policymakers. The results suggest that ratio-based models, particularly the Springate S-Score and Grover G-Score, can be utilized as more reliable tools for financial distress prediction and early warning systems. From an academic perspective, this study contributes to the literature by providing a more comprehensive evaluation of prediction models through the inclusion of precision and recall metrics, thereby offering a deeper understanding of model performance beyond accuracy alone.

However, this study is subject to several limitations. The sample is limited to 20 state-owned enterprises listed on the Indonesia Stock Exchange over the 2020–2023 period, which may restrict the generalizability of the findings to other sectors or types of firms. In addition, the study employs a single classification method, namely Naïve Bayes, without comparing its performance to other machine learning algorithms. Furthermore, although precision and recall have been incorporated, other evaluation metrics such as F1-score or cross-validation techniques were not explored in greater depth.

Future research is therefore recommended to expand the scope of analysis by including a larger and more diverse sample, such as private firms or companies from different industries and countries. Further studies may also compare multiple classification methods, including more advanced machine learning techniques, to assess the consistency of the results. In addition, integrating hybrid approaches that combine financial ratios and financial statement account data may provide a more balanced model in terms of accuracy and sensitivity. Such developments are expected to enhance the robustness and applicability of financial distress prediction models in both academic and practical contexts.

V. RECOMMENDATIONS

The findings of this research indicates that the Springate S-Score model exhibits the strongest accuracy in predicting corporate financial distress, future research is recommended to expand the dataset by including more companies and longer observation periods, as well as to apply and compare other machine learning methods such as decision trees, random forests, or support vector machines. Further studies may also explore the integration of additional financial and non-financial variables to enhance predictive performance and model robustness.

REFERENCES

[1] N. Deepa, J. Sathya Priya, and T. Devi, "Towards applying internet of things and machine learning for the risk prediction of COVID-19 in pandemic situation using Naive Bayes classifier for improving accuracy," in *Materials Today: Proceedings*, Elsevier Ltd, Jan. 2022, pp. 4795–4799. doi: 10.1016/j.matpr.2022.03.345.

[2] V. Sheth, U. Tripathi, and A. Sharma, "A Comparative Analysis of Machine Learning Algorithms for Classification Purpose," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 422–431. doi: 10.1016/j.procs.2022.12.044.

[3] O. Uludağ and A. Gürsoy, "On the Financial Situation Analysis with KNN and Naive Bayes Classification Algorithms," *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, vol. 10, no. 4, pp. 2881–2888, Dec. 2020, doi: 10.21597/jist.703004.

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

[4] T. N. Viet, H. Le Minh, L. C. Hieu, and T. H. Anh, "The naÏve bayes algorithm for learning data analytics," *Indian Journal of Computer Science and Engineering*, vol. 12, no. 4, pp. 1038–1043, Jul. 2021, doi: 10.21817/indjcse/2021/v12i4/211204191.

[5] S. Shetty, M. Musa, and X. Brédart, "Bankruptcy Prediction Using Machine Learning Techniques," *Journal of Risk and Financial Management*, vol. 15, no. 1, Jan. 2022, doi: 10.3390/jrfm15010035.

[6] X. Bredart, V. Vella, and J. Bonello, "Machine Learning Models For Predicting Financial Distress," *Journal of Research in Economics*, vol. 2, no. 2, pp. 174–185, Dec. 2018, doi: 10.24954/jore.2018.22.

[7] A. Malakauskas and A. Lakstutiene, "Financial distress prediction for small and medium enterprises using machine learning techniques," *Engineering Economics*, vol. 32, no. 1, pp. 4–14, Feb. 2021, doi: 10.5755/j01.ee.32.1.27382.

[8] E. I. Altman, "American Finance Association Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," 1968. [Online]. Available: http://www.jstor.orgURL:http://www.jstor.org/stable/2978933http://www.jstor.org/stable/2978933?seq=1&cid=pdf-reference#references_tab_contents

[9] R. Martanti, E. Lestari, M. Situmorang, M. Ilham, P. Pratama, and A. Talib Bon, "Financial Distress Analysis Using Altman (Z-Score), Springate (S-Score), Zmijewski (X-Score), and Grover (G-Score) Models in the Tourism, Hospitality and Restaurant Subsectors Listed on the Indonesia Stock Exchange Period," 2015.

[10] C. K. Andreou, P. C. Andreou, and N. Lambertides, "Financial distress risk and stock price crashes," *Journal of Corporate Finance*, vol. 67, Apr. 2021, doi: 10.1016/j.jcorpfin.2020.101870.

[11] R. Crespí-Cladera, A. Martín-Oliver, and B. Pascual-Fuster, "Financial distress in the hospitality industry during the Covid-19 disaster," *Tour. Manag.*, vol. 85, Aug. 2021, doi: 10.1016/j.tourman.2021.104301.

[12] R. Herdiana, R. Hmonangan, U. Hayati, and T. Suprapti, "Prediksi Financial Distress Perusahaan Food And Beverage Menggunakan Metode Naive Bayes," *Jurnal Ilmiah Betrik*, vol. 13, 2022.

[13] D. Wu, X. Ma, and D. L. Olson, "Financial distress prediction using integrated Z-score and multilayer perceptron neural networks," *Decis. Support Syst.*, vol. 159, Aug. 2022, doi: 10.1016/j.dss.2022.113814.

[14] N. Aydin, N. Sahin, M. Deveci, and D. Pamucar, "Prediction of financial distress of companies with artificial neural networks and decision trees models," *Machine Learning with Applications*, vol. 10, p. 100432, Dec. 2022, doi: 10.1016/j.mlwa.2022.100432.

[15] D. Dumitrescu, N. Bobitan, A. F. Popa, D. N. Sahlian, and C. A. Stanila, "Signaling Financial Distress Through Z-Scores and Corporate Governance Compliance Interplay: A Random Forest Approach," *Electronics (Switzerland)*, vol. 14, no. 11, Jun. 2025, doi: 10.3390/electronics14112151.

[16] G. L. Springate, "Predicting the Possibility of Failure in a Canadian Firm: A Discriminant Analysis.," Doctoral dissertation, Simon Fraser University, 1978.

[17] M. E. Zmijewski, "Methodological Issues Related to the Estimation of Financial Distress Prediction Models," 1984.

[18] J. Grover and A. Lavin, " Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy: a Service Industry Extension of Altman's Z-Score Model of Bankruptcy Prediction. ," 2001.

[19] D. J. Maulana, S. Saadah, and P. E. Yunanto, "Kmeans-SMOTE Integration for Handling Imbalance Data in Classifying Financial Distress Companies using SVM and Naïve Bayes," *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 10, no. 1, pp. 54–61, 2024, doi: 10.29207/resti.v8i1.5150.

[20] S. R. Aadilah and T. P. Hadi, "Analisis Tingkat Kebangkrutan Sebelum dan Saat Pandemi Dengan Menggunakan Metode Altman Z-Score dan Springate S-Score," *Dinamika Akuntansi, Keuangan dan Perbankan*, vol. 11, no. 1, 2022.

[21] P. Saha and S. Ahmed, "Comprehensive Analysis of Altman Z Score, Zmijewski X Score, Springate S-Score and Grover G-Score Model for Predicting Financial Health of Listed Non-Bank Financial Institutions (NBFIs) of Bangladesh," *Open Journal of Business and Management*, vol. 12, no. 05, pp. 3342–3365, 2024, doi: 10.4236/ojbm.2024.125167.

[22] M. Elhoseny, N. Metawa, G. Sztano, and I. M. El-hasnony, "Deep Learning-Based Model for Financial Distress Prediction," *Ann. Oper. Res.*, vol. 345, no. 2, pp. 885–907, Feb. 2025, doi: 10.1007/s10479-022-04766-5.

[23] J. P. Noriega, L. A. Rivera, and J. A. Herrera, "Machine Learning for Credit Risk Prediction: A Systematic Literature Review," *Data (Basel).*, vol. 8, no. 11, Nov. 2023, doi: 10.3390/data8110169.

Received: Jan 13th, 2026 | Published: March 30th, 2026