Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI

(Rekayasa Sistem dan Teknologi Informasi)

Vol. 9 No. 3 (2025) 564 - 569 e-ISSN: 2580-0760

Prediction of Financial Distress in Retail Companies Using Long-Short Term Memory (LSTM)

Wahyuni Windasari[1*], Tuti Zakiyah[2]

[1]Department of Data Science, Faculty of Science and Technology, Universitas Putra Bangsa, Kebumen, Indonesia

[2]Department of Management, Faculty of Economics and Business, Universitas Putra Bangsa, Kebumen, Indonesia

[1]wahyuwindasari@gmail.com, [2]tutizakiyah@gmail.com

Abstract

Financial distress is a condition in which an entity struggles to meet its debt and operating obligations.. Financial distress can lead to bankruptcy or company closure if corrective action is not taken. This study aims to forecast financial distress in retail companies by utilizing key financial ratios, including Total Asset Turnover (TATO), Current Ratio (CR), Return on Assets (ROA), and Debt-to-Equity Ratio (DER). The analysis is based on secondary data from Indonesian retail companies listed on the Indonesia Stock Exchange (IDX) during the 2022–2024 period. The dataset exhibited missing values and class imbalance, which were addressed using mean imputation and the Synthetic Minority Oversampling Technique (SMOTE), respectivelyTo perform predictions, a Long Short-Term Memory (LSTM) model was implemented. The integration of SMOTE contributed to enhanced detection of the minority class; however, it was accompanied by a slight reduction in overall predictive accuracy. The model demonstrated a performance accuracy of 86%, with a recall rate of 85%, a precision of 100%, and an F1-score of 92%.

Keywords: deep learning; LSTM; prediction; financial distress; retail

How to Cite: W. Windasari and T. Zakiyah, "Prediction of Financial Distress in Retail Companies Using Long-Short Term Memory (LSTM)", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 9, no. 3, pp. 564 - 569, Jun. 2025.
*Permalink/*DOI: https://doi.org/10.29207/resti.v9i3.6217

Received: December 7, 2024
Accepted: June 3, 2025
Available Online: June 18, 2025

1. Introduction

Financial distress represents a stage of financial deterioration that a company experiences prior to bankruptcy [1]. It is characterized by the company's inability to generate sufficient profits or a tendency to incur ongoing losses [2]. Although financial distress does not necessarily lead to bankruptcy [3], a prolonged and significant decline Poor financial performance can elevate the risk of corporate failure, potentially leading to bankruptcy, which affects many firms, may negatively impact the global economy and harm various stakeholders involved [4].

The information on financial distress is beneficial for financial reporting managers and stakeholders in developing strategies to avoid bankruptcy [5]. Therefore, developing financial distress prediction models using various explanatory variables and statistical methods is essential to provide insights into a company's financial condition [6]. Financial distress prediction has attracted significant attention from both academics and practitioners, with numerous models proposed over the past decades [7]. Financial distress can be predicted by analyzing a company's financial ratios [8]. These typically reflect a firm's capacity to manage working capital, generate sales, handle debt, and derive profits from both sales and asset utilization [9].

A review of financial distress studies in Indonesia indicates that commonly used prediction models include the Springate, Altman Z-Score, and Zmijewski methods [10]. The financial distress model, especially the Altman model, has limitations, namely that this model comes from empirical studies on manufacturing companies, so that when this model is used in companies other than manufacturing, there are many irregularities where the prediction results given are almost 100% bankruptcy in the sample studied or inversely proportional to the actual conditions [11]. Prediction errors can also be caused because the model uses time series data from financial ratios that are susceptible to missing values and class imbalances. Therefore, the development of predictive analysis models using time series data continues to be carried out. One of them uses machine learning and deep learning model approaches. This model allows for learning with more complex layers to obtain high accuracy and is more efficient in providing more accurate forecasts and modeling linear and nonlinear data compared to traditional forecasting methods. Research related to the comparison of financial distress models provides empirical analysis results that show that the Perceptron Neural Network Multilayer model has a better level of precision in forecasting Chinese financial data in comparison to the Z-Score statistical model [12].

The results of this comparison show that LSTM has a higher accuracy performance value compared to RNN. Research on financial distress prediction in Indonesia remains limited, particularly regarding the application and optimization of deep learning models [13]. Employing deep learning approaches is essential to develop more robust predictive models. Among these, Long Short-Term Memory (LSTM) has demonstrated strong performance in time series analysis [14]. LSTM, a time series–based algorithm, is widely recognized for its reliability and accuracy in forecasting tasks. The choice of LSTM in this study is supported by comparative research showing that LSTM outperforms standard Recurrent Neural Networks (RNN) in terms of prediction accuracy [15].

Based on this background, this study will conduct an analysis related to financial distress in Indonesia using the deep learning method based on LSTM by considering the handling of missing values and class imbalances in the data used. This study contributes to advancing financial distress prediction in Indonesia by proposing a high-accuracy model tailored for the retail sub-sector.

2. Methods

This research uses secondary data comprising financial ratios number of Indonesian retail businesses that are listed on the Indonesia Stock Exchange (IDX) from 2022 to 2024. The selected companies consistently published quarterly financial statements during this period and provided complete ratio data covering liquidity, solvency, activity, and profitability indicators.

A company's capacity to satisfy short-term obligations with its present assets is gauged by the liquidity ratio, and in this research current ratio (CR) serves as a stand-in for it. The ratio of solvency, represented by the Debt to Equity Ratio (DER), compares a company’s total debt to its equity; a high solvency ratio may indicate potential difficulty in meeting debt obligations. The activity or performance ratio, proxied by the Total Asset Turnover (TATO), evaluates the company’s operational efficiency, including its ability to sell products, collect receivables, and manage asset utilization. Return on Assets (ROA) is used to evaluate profitability, which indicates a firm’s efficiency in generating earnings from its assets [16].

Financial distress is indicated by the Debt-Service Coverage Ratio (DSCR), It assesses a company's capacity to pay off existing debt with cash flow [17]. The DSCR is selected as a proxy for financial distress because it reflects a company’s inability or delay in meeting its short-term and long-term debt obligations [18]. In addition to being a helpful gauge of a business's debt capacity, the DSCR can help guide choices about complementing equity investments [19]. A DSCR value of ≤ 1.2 is typically interpreted as an indication of financial distress [20].

This study adopts a deep learning approach utilizing One kind of artificial neural network (ANN) is called long short-term memory (LSTM) capable of learning complex data patterns through multiple layers. The model identifies hidden patterns in the data and classifies them to generate accurate outputs for new inputs [21].

The following is a figure of the LSTM-based deep learning architecture.

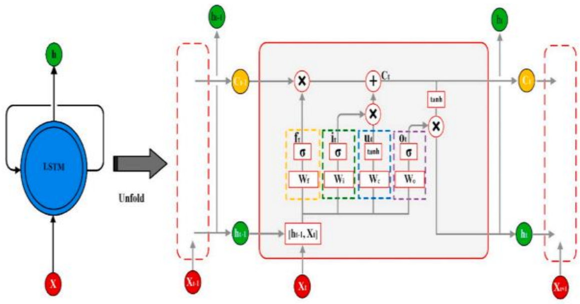

Figure 1. LSTM Architecture [22]

According to Figure 1, The forget, input, and output gates are the three primary gates that make up the LSTM architecture. Memory reset gate (f_t) has a function to delete information from the cell status. The forget gate is responsible for identifying which components The decision to keep or erase from the previous cell state Ct−1. This process is mathematically represented in Equation 1.

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \tag{1}$$

The input gate (i_t) has functions to decide on new information entering the cell state. This section consists of two parts: the gate vector i_t, calculated using the sigmoid function, and the candidate cell state $\overline{C_t}$ calculated using the tanh activation function. These two parts are expressed in The second and third equations.

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \tag{2}$$

$$\overline{C_t} = tanh(w_c[h_{t-1}, x_t] + b_c) \tag{3}$$

Next, the stored information from the forget gate is combined with the newly selected information from the input gate to update the State of cell C_t, as shown in Equation 4.

$$C_t = ft * C_{t-1} + i_t * \overline{C}_t \quad (4)$$

Finally, the output gate ($\mathrm{i_t}$) functions to filter relevant Utilize the current cell state to compute and deliver the output, as expressed in Equation 5.

$$h_t = Ot * \tanh C_t \quad (5)$$

The effectiveness of the LSTM model is highly dependent on its architecture. The hyperparameters influencing LSTM performance include learning rate, batch size, and number of epochs [23]. The parameters to be used in this study are proposed as in Table 1.

Table 1. LSTM Parameter Description

Parameter	Value
Learning rate	0.01 , 0.001
Epoch	5, 10, 20
Batch	16, 32, 64, 128

Performance evaluation is an essential stage in the life cycle of deep learning models. There are two commonly used techniques for classification model performance, namely confusion metrics and classification report [21].

A confusion metrics is a sized table (with the number of classes/labels/categories) that contains information about the correct or incorrect number of predictions of a classification model, which helps compare actual values with predicted values. The real class is represented by each row in the confusion metrics, whereas the projected class is shown by each column.

The confusion metrics produces four types of values consisting of True Positive (TP): Positive predictions & positive actual values, True Negative (TN): Negative predictions & negative actual values, False Positive (FP): Positive predictions & negative actual values, and False Negative (FN): Negative predictions & positive actual values.

The classification report is one of the measures of model performance, and calculations are obtained from the confusion metrics. Model performance metrics include F-1 score, recall, accuracy, and precision.

The accuracy describes how accurately the model classifies correctly, is defined in Equation 6.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (6)$$

The precision describes the accuracy between the requested data and the model's prediction results, is shown in Equation 7.

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

The recall describes the model's success in rediscovering information, is given in Equation 8.

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

The F-1 score describes the average ratio of precision and weighted recalls, is defined in Equation 9.

$$F1 - Score = 2\ x\ \frac{Recall\ x\ Precision}{Recall+Precision} \quad (9)$$

3. Results and Discussions

This research utilizes secondary data sourced from the Indonesia Stock Exchange (IDX), comprising 216 records categorized into 24 financially distressed and 192 financially healthy companies. The data exhibits a substantial class imbalance, with 88.89% of the records labeled as healthy (label 0) and only 11.11% as distressed (label 1). This imbalance presents a challenge for model training and evaluation. Furthermore, the dataset contains missing values that must be addressed to ensure accurate analysis and robust model performance. The variation in the DSCR over time ratio of the entire retail sub-sector during the research observation period is given in the Figure 2.

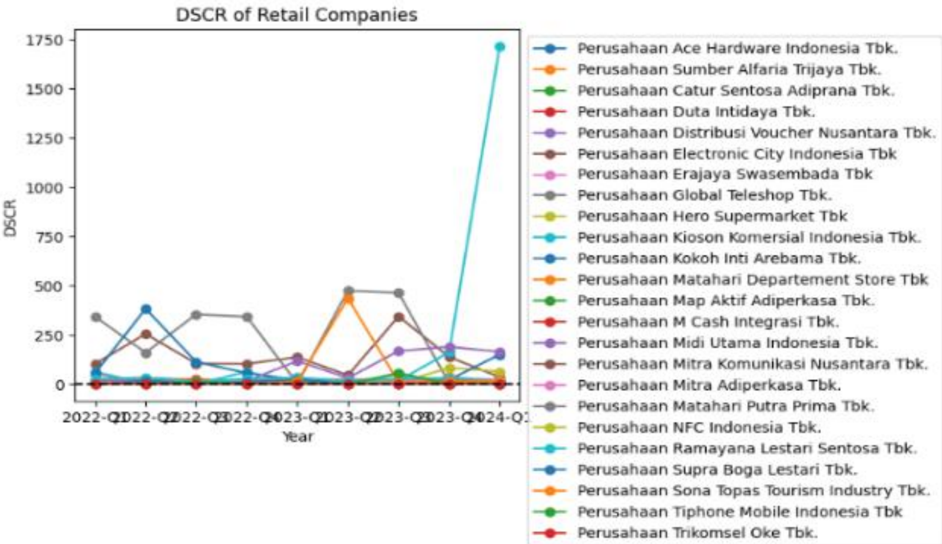

Figure 2. DSCR of Retail Sub-Companies and Financial Distress Threshold

From Figure 2, it can be seen that several companies have a DSCR value of ≤ 1.2, especially the initial period of the study, namely the observation period of the First Quarter of 2022. This indicates that many retail sub-companies are experiencing financial distress. This study utilizes financial report tables as its research instrument, which include CR (Current Rasio/ X_1), DER (Debt to Equity Ratio/ X_2), TATO (Total Assets Turnover / X_3), and ROA (Return On Aset /X_4) which will be used to explain DSCR [24].

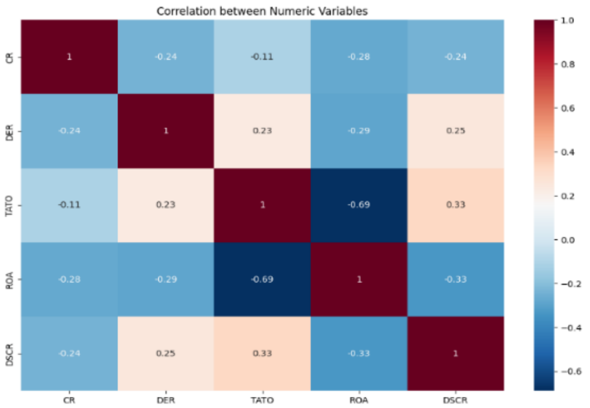

Figure 3. Heatmap of Correlation Metrics Between Variables

Figure 3 shows how the relationship between variables in the study. From the heatmap image, it can be seen that the ROA and TATO variables have the strongest negative correlation among the others. While the relationship between CR and TATO has the weakest negative correlation among the others. Although the overall correlation between the independent variables and the DSCR is low, this is visually represented by the bright color shades in each corresponding box.

In the data processing stage, the collected dataset is first examined for missing values, duplicate entries, and other data errors before being input into the model. The data preprocessing steps in this study include normalization, handling missing financial data using interpolation and mean imputation methods [25], and mitigating class imbalance using the Synthetic Minority Oversampling Technique (SMOTE) [26].

The data in this study contains 8 missing values, namely CR, DER, TATO, and ROA data for the 2022 period, Quarter 2 and Quarter 3. Handling of missing data using the interpolation and mean methods. The MSE error results from both methods are shown in Figure 4.

Based on Figure 4, the smallest MSE value is obtained from handling missing values using the mean method. Furthermore, the data that will be used for LSTM analysis is the imputed data using the mean method.

Upon completion of the data preprocessing stage, The dataset is divided into training and testing subsets, and the proportion of this split plays a crucial role in determining the model's accuracy and overall performance. To identify the optimal configuration, this study evaluates two scenarios with test-to-training data ratios of 10:90 and 20:80. During the training phase, the model is fitted using the training dataset. In the subsequent testing phase, the model's performance and predictive capability are evaluated using the test dataset.

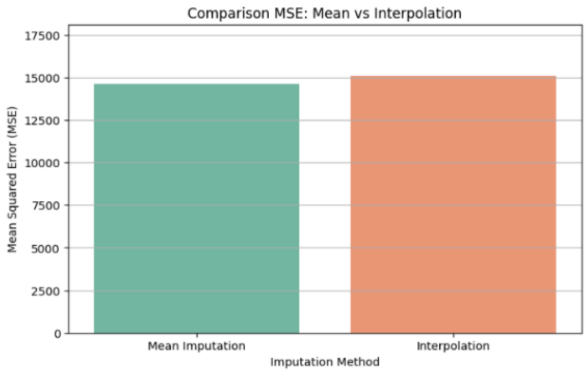

Figure 4. Comparison of MSE Values of Handling Missing Value

The model development process begins with selecting appropriate hyperparameters, which often lack definitive guidelines and are therefore adjusted according to the researcher's requirements through a trial-and-error approach. The combination of experimental variables employed in this study is depicted in Figure 5.

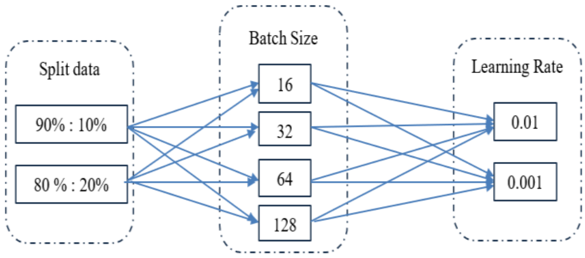

Figure 5. Combination of LSTM Model Experimental Parameters

The training process is carried out using experimental variables combined with split data, batch size, and learning rate (∝). The selection of parameters is based on the results of experiments conducted during this study, to achieve the highest model performance. The summary results of the selected parameters that optimize the LSTM model training process more efficiently are listed in Table 2.

Table 2. Training Parameters

Parameter	Value
Loss Function	Categorical Crossentropy
Optimizer	Adam
Learning Rate	0.001
Batch Size	16
Epoch	20

After selecting the appropriate hyperparameters, To evaluate its effectiveness in classifying imbalanced data, the performance of the LSTM model is measured using confusion matrix metrics. These indicators deliver an in-depth analysis of the model's performance in correctly identifying both positive and negative classes. The results of the model testing, both prior to and following the implementation of SMOTE, are displayed in Table 3.

Table 3. Model Performance

Algorithm	Accuracy	Recall	Precision	F1-Score
Non-SMOTE	0.93	1.00	0.93	0.96
SMOTE	0.86	0.85	1.00	0.92

The test results indicate that the application of SMOTE leads to a decrease in accuracy, recall, and F1-score. Despite these changes, the LSTM model with SMOTE still demonstrates strong performance, as evidenced by a higher precision compared to the model without SMOTE and an F1-score of 92%, which is comparable to the score before applying SMOTE. This suggests that while SMOTE reduces accuracy, it positively enhances the model's ability to handle class imbalance without significantly compromising its overall classification capability. The confusion metrics results for the test data are presented in Figure 6.

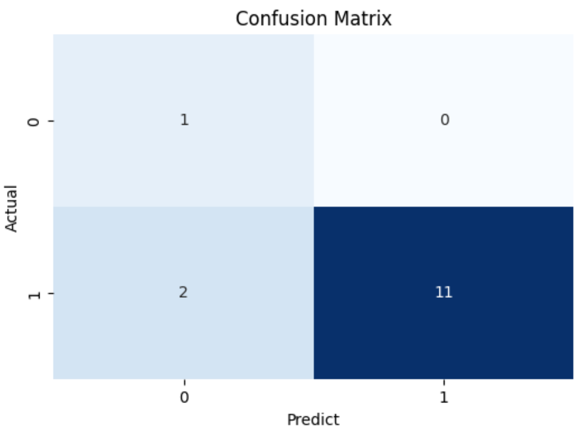

Figure 6. Confusion Metrics Data Testing

The confusion metrics results of the LSTM model, shown in Figure 6, provide detailed insights into the model's predictions for both the negative and positive classes. The model demonstrates strong performance, as indicated by the high number of correctly predicted positive cases.

4. Conclusions

This study concludes that handling missing values using the mean imputation method, combined with addressing class imbalance through SMOTE in the LSTM model, yields strong predictive performance, achieving an accuracy of 86%, a precision of 100%, a recall of 85%, and an F1 score of 92%. These results suggest that SMOTE effectively balances class distribution without significantly compromising classification ability. The accurate predictions generated by this LSTM model can assist investors in making informed investment decisions and enable retail companies to implement early risk mitigation strategies, such as enhancing liquidity, optimizing cash flow management, and strengthening capital reserves. It is advised that company management and investors diligently monitor financial ratios, encompassing liquidity, solvency, activity, and profitability metrics, as they have been demonstrated to be reliable indicators for predicting financial distress.

References

[1] A. A. Lestari and D. E. Kusrini, "Analisis Financial Distress Menggunakan Regresi Data Panel," Semin. Nas. Off. Stat., vol. 2021, no. 1, pp. 264–273, 2021, doi: 10.34123/semnasoffstat.v2021i1.855.

[2] A. Larasati and A. Jayanih, "The Influence of Company Size, Company Growth and Profitability on Financial Distress," Scientia, vol. 2, no. 1, pp. 16–23, 2023, doi: 10.51773/sssh.v2i1.128.

[3] N. B. M. Maghrebi, "Preliminary Evidence on the Relation between Economic Uncertainty, Financial Stress and Liquidity Risk," Wakayama Econ. Rev., vol. 408, pp. 17–34, 2022, [Online]. Available: https://repository.center.wakayama-u.ac.jp/4833

[4] P. R. Masdiantini and N. M. S. Warasniasih, "Laporan Keuangan dan Prediksi Kebangkrutan Perusahaan," J. Ilm. Akunt., vol. 5, no. 1, p. 196, 2020, doi: 10.23887/jia.v5i1.25119.

[5] R. M. E. Lestari, M. Situmorang, M. I. P. Pratama, and A. T. 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 2015-2019," Proc. Int. Conf. Ind. Eng. Oper. Manag., pp. 4249–4259, 2021, doi: 10.46254/an11.20210755.

[6] N. Bărbuță-Mișu and M. Madaleno, "Assessment of Bankruptcy Risk of Large Companies: European Countries Evolution Analysis," J. Risk Financ. Manag., vol. 13, no. 3, 2020, doi: 10.3390/jrfm13030058.

[7] A. Mehmood and F. De Luca, "Financial distress prediction in private firms: developing a model for troubled debt restructuring," J. Appl. Account. Res., 2023, doi: 10.1108/JAAR-12-2022-0325.

[8] H. Sasongko, A. Fajar Ilmiyono, and A. Tiaranti, "Financial Ratios and Financial Distress in Retail Trade Sector Companies," JIAFE (Jurnal Ilm. Akunt. Fak. Ekon., vol. 07, no. 01, pp. 63–72, 2021, doi: 10.34204/jiafe.v7i1.3380.

[9] T. Zakiyah and W. Windasari, "Implementasi Model Finansial Distres Pada Perusahaan Manufaktur Yang Terdaftar di IDX Tahun 2015-2017," J. Ilm. Akunt. dan Keuang., vol. 9, no. 1, pp. 61–74, Jan. 2020, doi: 10.32639/jiak.v9i1.330.

[10] D. Nurfauziyyah and A. I. Muslim, "Literature Review tentang Financial Distress yang Terbit di Jurnal Sinta," vol. 9, no. 02, pp. 225–240, 2024.

[11] M. N. Syaifudin, "Perbandingan Analisis Prediksi Kebangkrutan Menggunakan Model Altman dengan Ukuran Perusahaan Sebagai Variabel Penjelas pada Perusahaan Industri Keuangan Bank yang Terdaftar di Bursa Efek Indonesia Indonesia Periode 2010-2012," J. Mhs. Univ. Negeri Surabaya, pp. 1–25, 2012.

[12] 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, no. March, p. 113814, 2022, doi: 10.1016/j.dss.2022.113814.

[13] M. El-Bannany, M. Sreedharan, and A. M. Khedr, "A robust deep learning model for financial distress prediction," Int. J. Adv. Comput. Sci. Appl., no. 2, pp. 170–175, 2020, doi: 10.14569/ijacsa.2020.0110222.

[14] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, "A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU," no. September, 2023, doi: 10.13140/RG.2.2.11938.81609.

[15] S. S. Nurashila, F. Hamami, and T. F. Kusumasari, "Perbandingan Kinerja Algoritma Recurrent Neural Network (Rnn) Dan Long Short-Term Memory (Lstm): Studi Kasus Prediksi Kemacetan Lalu Lintas Jaringan Pt Xyz," JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform., vol. 8, no. 3, pp. 864–877, 2023, doi: 10.29100/jipi.v8i3.3961.

[16] M. Laksmiwati, A. Mappadang, A. Indrabudiman, and V. G. Riza, "Determination of Financial Performance on Prediction Financial Distress," Asian J. Econ. Bus. Account., no. December 2021, pp. 46–57, 2021, doi: 10.9734/ajeba/2021/v21i2230523.

[17] A. Rabbani, N. A. Achsani, and T. Andati, "Financial Distress Analysis of Delisted Companies on Indonesian Stock Exchange for Period of 2013-2019," Russ. J. Agric. Socio-Economic Sci., vol. 119, no. 11, pp. 11–18, 2021, doi: 10.18551/rjoas.2021-11.02.

[18] G. Difoasih and P. Sihombing, "Analysis of The Effect of Financial Performance, Board Gender Diversity and CSR on Financial Distress in Tourism on the IDX," Int. J. Innov. Sci. Res. Technol., vol. 8, no. 12, 2023.

[19] M. Marcellino, G. Castelblanco, and A. De Marco, "Multiple linear regression model for project's risk profile and DSCR," AIP Conf. Proc., vol. 2928, no. 1, 2023, doi: 10.1063/5.0171033.

[20] J. Ruster, "Mitigating Commercial Risks in Project Finance. Public Policy for Private Sector," in Public Policy for Private Sector, New York, 1996.

[21] M. R. R. Allam and A. T. Wibowo, "Klasifikasi Genus Tanaman Anggrek berdasarkan Citra Kuntum Bunga Menggunakan Metode Convolutional Neural Network (CNN)," e-Proceeding Eng., vol. 7, no. 2, pp. 8015–8023, 2020, [Online]. Available: https://www.programmersought.com/article/3724355693/

[22] S. Ghimire, R. C. Deo, D. Casillas-Pérez, S. Salcedo-Sanz, E. Sharma, and M. Ali, "Deep learning CNN-LSTM-MLP hybrid fusion model for feature optimizations and daily solar radiation prediction," Meas. J. Int. Meas. Confed., vol. 202, no. April, p. 111759, 2022, doi: 10.1016/j.measurement.2022.111759.

[23] A. Shafi, S., Assad, "Exploring the Relationship Between Learning Rate, Batch Size, and Epochs in Deep Learning: An Experimental Study," in Soft Computing for Problem Solving, 2023, pp. 201–209. doi: https://doi.org/10.1007/978-981-19-6525-8_16.

[24] D. F. Santosa, L. Anggraeni, and K. Pranowo, "Determinan Financial Distress Perusahaan Subsektor Ritel Di Bursa Efek Indonesia," J. Apl. Bisnis dan Manaj., vol. 6, no. 1, pp. 128–141, 2020, doi: 10.17358/jabm.6.1.128.

[25] S. J. Pipin, R. Purba, and H. Kurniawan, "Prediksi Saham Menggunakan Recurrent Neural Network (RNN-LSTM) dengan Optimasi Adaptive Moment Estimation," J. Comput. Syst. Informatics, vol. 4, no. 4, pp. 806–815, 2023, doi: 10.47065/josyc.v4i4.4014.

[26] N. Istiana and A. Mustafiril, "Perbandingan Metode Klasifikasi pada Data dengan Imbalance Class dan Missing Value," J. Inform., vol. 10, no. 2, pp. 101–108, 2023, doi: 10.31294/inf.v10i2.15540.