p–ISSN: 2723 - 6609 e-ISSN: 2745-5254
Vol. 5, No. 8 August 2024 http://jist.publikasiindonesia.id/

Machine Learning Analysis in Predicting Bankruptcy in Companies (Case Study of Manufacturing Companies Listed on the Stock Exchange)

Citra Yustika Pratiwi[1*], Siti Nurwahyuningsih Harahap[2]
Universitas Indonesia, Indonesia
Email: citrayustika08@gmail.com
*Correspondence

Keywords: bankruptcy, machine learning, manufacturing companies.

ABSTRACT

This study aims to analyze bankruptcy prediction for manufacturing companies using machine learning. Financial data from manufacturing companies listed on the Indonesia Stock Exchange for the period from 2013 to 2023 are used in this study. The analytical methods employed include Long Short-Term Memory (LSTM), Support Vector Machine (SVM), Random Forest, and Extreme Gradient Boosting (XGBoost). The results of this study are expected to provide benefits to various stakeholders: manufacturing companies in identifying early signs of bankruptcy, creditors in evaluating the feasibility of extending credit, investors in making investment decisions, academics in advancing research in bankruptcy prediction, and market regulators (OJK) in enhancing the efficiency of supervision over manufacturing companies. The results indicate that SVM is effective in predicting historical data with consistent performance, while LSTM excels in handling variations and patterns in new data.

Introduction

The manufacturing industry has a crucial role in the Indonesian economy, as evidenced by its significant contribution to Gross Domestic Product (GDP) since the 1980s (Madjid, Mahdi, Lukito, Nofri, & Prasvita, 2021). This sector continues to develop rapidly, showing stable growth with GDP in the manufacturing sector in 2021 reaching IDR 2,946.9 trillion and investment reaching IDR 325.4 trillion, as well as being a source of employment for 1.2 million new people (Ministry of Industry, 2022). Indicators such as the Purchasing Managers Index (PMI) also recorded record highs, reflecting the sector's strong expansion and its role as a key pillar in national economic growth (Joshi, Ramesh, & Tahsildar, 2018).

Even though the manufacturing industry shows positive growth, economic challenges remain an important factor influencing the performance of companies in this sector. Economic fluctuations can trigger financial difficulties, which is a critical phase before the risk of bankruptcy (Swari & Pristiana, 2020). This phenomenon, known as financial distress, is characterized by decreased income, negative cash flow, and increased debt that can threaten long-term business continuity (Siswoyo, 2020).

Bankruptcy prediction is crucial in managing a company's financial risk. By applying machine learning techniques such as the Altman Model and Ohlson Model, companies can identify and manage risks more effectively (Muta'ali, 2019). This model uses historical financial data to produce accurate bankruptcy scores, assisting companies in making strategic decisions to maintain financial stability and business sustainability (Shetty & Kellarai, 2022).

(Kothuru et al., 2022), this study suggests that Random Forest is effective in handling large and complex datasets and provides estimates of the importance of variables in bankruptcy prediction. They suggest evaluating traditional models with various machine learning techniques to provide a more comprehensive and relevant picture. (Sulastri, 2014), they compared the Ohlson and Altman models in bankruptcy prediction, with Altman proving to be more effective in the context of bankruptcy prediction for large and small companies. This study suggests combining traditional models with machine learning algorithms as well as evaluation with various metrics to provide a more in-depth picture (Almas, 2023).

Based on the background above, the main objective of this research is to evaluate machine learning models that can produce the best bankruptcy predictions and models that have the highest prediction accuracy.

Research Methods

This research uses an archival study research strategy with a focus on quantitative comparative analysis. The method applied is predictive analysis using financial report data from manufacturing companies listed on the Indonesia Stock Exchange (BEI). The main data is obtained from financial reports submitted by these companies via the official IDX website. This research selected companies that have published annual reports from 2013 to 2023 as samples, using a purposive sampling method to ensure relevance to the research objectives. The variables analyzed include various financial ratios adopted from the Altman and Ohlson model to predict potential bankruptcy. Data analysis was carried out through a preprocessing process which included removing outliers using a Z-score, dividing the dataset into training and validation data with a ratio of 80:20, as well as feature scaling using StandardScaler to ensure variable scale consistency. The creation of a machine learning model is based on the reputation and effectiveness of the Altman and Ohlson model in predicting corporate bankruptcy. This analysis aims to produce an accurate predictive model to support decision making regarding financial risk management of manufacturing companies in Indonesia.

Research Approach

This study uses a quantitative methodology as a framework for comparative analysis. (Creswell et al., 2018) define quantitative research as a research method that tests theory by measuring variables and analyzing numerical data using statistical procedures. This approach aims to determine relationships between variables, test hypotheses, and make predictions. The quantitative approach in this research is in order to obtain an in-depth and comprehensive understanding of the use of machine learning in predicting bankruptcy in manufacturing companies.

The technique applied in this study is predictive analysis. Predictive analytics is a data analysis technique used to predict future outcomes based on historical data (Qi & Tao, 2018). Predictive analytics can be used to predict corporate bankruptcy, so that companies can take preventative action and strategic adjustments before experiencing significant financial difficulties

Data Source

The information used in this study is sourced from shortage reports of manufacturing companies that are registered on the Indonesian Stock Exchange (BErI). The main data is obtained from shortfall reports submitted by terrsburt companies which can be accessed through the official BERI website. Apart from that, other relevant data is GNP (Gross National Product) which can be obtained from trusted sources such as financial institutions, government institutions or economic research institutions.

Sample Determination Method

The sample in this research consists of manufacturing companies that are registered with BERI and have published financial statements in the time period 2013 to 2023.The sample selection process is carried out by using a purposive sampling method. This method selection allows selecting samples that are relevant to the research objectives. The criteria for sample selection are various:

1. Manufacturing companies that are registered with Burrsa Erferk Indonesia and have published financial reports for the period 2013 to 2023.
2. The company has complete data regarding the relevant variables used in the research.

Research Variables

The variables analyzed in this study are divided into two types, namely independent variables and dependent variables. Independent variables include financial ratios such as liquidity, profitability, solvency, activity and market dimensions. Meanwhile, the dependent variable is the company's financial health status, which is represented by a binary variable where the number 1 indicates bankruptcy and 0 indicates non-bankruptcy.

Results and Discussion

The population in this study is manufacturing companies listed on the Indonesia Stock Exchange (IDX) in 2013-2023. This study uses two types of datasets, namely the Ohlson and Altman Z model datasets. Each dataset has different attributes because it adapts its respective model and predefined labels to the model's calculations.

The dataset used in this study is the financial report data of 161 manufacturing companies which is secondary data obtained from data sources on the www.idx.co.id website.

Table 1
Research Sample Selection Procedure

It	Criterion	Sum

1	Manufacturing Companies listed on the IDX in the period 2013-2023	166
2	Companies that have not submitted financial statements	5
	Total research observations	161

Table 1 shows that during the period 2013-2023, there were a total of 166 manufacturing companies. Of these, 5 companies did not publish financial statements during the period. Thus, 161 banking companies meet the sample criteria for this study. Furthermore, companies that meet the sample criteria are grouped into two categories: Category 1 for companies that are experiencing financial distress or bankruptcy, and Category 0 for companies that are not experiencing financial distress or not bankrupt (Ariyanto, 2017).

Model Formation

The process of forming a classification model aims to create a classification model. The model will be used to classify the labels for both datasets. The model formation process uses the scikit-learn library and the Python programming language. 4 models will be formed in this study, including Support Vector Machine (SVM), Random Forest, XGBoost, and long short term memory (LSTM).

Support Vector Machine (SVM)

The SVM model formation process uses hyperparameter tuning techniques to determine the best parameters to be used on the model. This technique uses the Grid Search CV function derived from the scikit-learn library in the python programming language. For each model training process with training data with certain parameters, the model will be evaluated with K-Fold cross-validation with a cv value equal to 5 (Wibowo, 2012).

The Support Vector Machine (SVM) is divided into two different datasets, namely Ohlson data and Altman data. In the first part, SVM is applied to Ohlson data with hyperparameter settings through Grid Search Cross-Validation. After getting the best model, predictions are made on the test data and calculation of evaluation metrics such as accuracy, precision, recall, F1 score and specificity.

Random Forest

The Random Forest model formation process uses hyperparameter tuning techniques to determine the best parameters to be used in the model. This technique uses the GridSearchCV() function derived from the scikit-learn¬ library in the Python programming language. For each process of training a model with training data with certain parameters, the model will be evaluated with K-Fold cross-validation with cv=5. The modelling uses the Random Forest algorithm with a variety of predefined parameters. First, the best parameter search was carried out using the GridSearchCV method with cross-validation 5 times. The best results of the model along with the parameters used and the best score are displayed. Then, predictions were made on the test dataset using the best model obtained, followed by the calculation and printing of evaluation metrics such as precision, recall, specificity and F1-score to evaluate the model's performance on the test data. This process is repeated for two different data sets, "Houston" and "Saltzman", with the same steps.

XGBoost

The XGBoost model formation process uses hyperparameter tuning techniques to determine the best parameters to be used on the model. This technique uses the GridSearchCV() function derived from the scikit-learn¬ library in the Python programming language. For each model training process with training data with certain parameters, the model will be evaluated with K-Fold cross-validation with a value of cv=5.

GridSearchCV along with XGBClassifier is used to optimize key parameters such as max_depth, learning_rate, and subsamples to improve the accuracy of the classification model. param_grid explicitly defines a range of values for each parameter, which allows XGBClassifier to be tested in a variety of configurations through cross-validation five times by GridSearchCV.

Long short term Memory (LSTM)

Data training needs to be reshaped to change the dimensions before forming the LSTM model. The model has an epoch parameter of 20 and a hidden_units of 64.

Each model is arranged sequentially with an LSTM layer that has 64 units, followed by sigmoid activation and a Dense layer. The data is rearranged to meet the LSTM input format, and the model is compiled with the Adam optimizer and the mean squared error loss function. The training was carried out for 20 epochs.

Model Analysis

Model analysis is carried out to obtain a classification model with parameters that have the highest accuracy value. The model analysis will be carried out on both datasets. Table 2 is a comparison of the accuracy values of the classification model along with the best parameters.

Table 2
Comparison of Classification Models

Dataset	Model	*Accuracy* (%)	*Precision* (%)	*Recall* (%)	*F1-Score* (%)	*Specificity* (%)
Ohlson	LSTM	91.79	82.35	43.75	57.14	98.66
	SVM	99.21	96.87	96.87	96.87	99.55
	Random Forest	98.04	93.54	90.62	92.06	99.10
	XGBoost	97.65	88.23	93.75	90.90	98.21
Altman	LSTM	93.86	89.65	91.76	90.69	94.88
	SVM	98.85	100	96.47	98.20	100
	Random Forest	94.63	93.82	89.41	91.56	97.15
	XGBoost	96.16	97.31	88.23	93.75	100

Based on the results from the performance table, there are five algorithm options to consider:

1. High Accuracy: Choose a machine learning technique with high accuracy4 if the most important thing is how accurate the system is in classifying data correctly. Accuracy is the ratio of correct predictions (both positive and negative) to the accuracy of the

data. From the table, it can be seen that the machine learning technique with the highest accuracy in Modern Ohlson and Modern Altman is SVM.

2. High Recall: Choose a machine learning technique with high recall if the error calculation is more likely to cause Falser Positive than Falser Nergative. In this study, it is better for the model to incorrectly predict a company that is actually not bankrupt as bankrupt than to incorrectly predict a company that is actually bankrupt as not bankrupt. From the table, it can be seen that the machine learning technique with the highest frequency of calls on Model Ohlson and Model Altman is SVM.
3. High Precision: Choose a machine learning technique with high precision if you prefer to take truer positives and avoid false positives. In this study, it is better for the model to incorrectly predict a bankrupt company that is not actually bankrupt than to incorrectly predict a non-bankrupt company that is actually bankrupt. From the table, it can be seen that the algorithm with the highest precision in Model Ohlson and Model Altman is SVM.
4. High Specificity: Choose a machine learning technique with high specificity if taking errors does not really want a Falser Positive to occur. The model should avoid falsely detecting bankruptcy in companies that are not actually bankrupt. From the table, it can be seen that the algorithm with the highest specificity in Model Ohlson is Random Forest and Model Altman is SVM.
5. High F1 Scorer: Choose a machine learning technique with high F1 Scorer if the calculation of the error is more concerned with the balance between recall and precision. This means that the chosen algorithm must have small Falser Positive and Falser Negative values. From the table, it can be seen that the highest recall algorithm in Model Ohlson and Model Altman is SVM

Taking into account the metrics that best suit the distress analysis needs, SVM appears to be a consistent and superior choice for the most important metrics based on the results of the performance table.

Performance information is presented in numerical form only. To display the performance information of the classification algorithm graphically, the Receiver Operating Characteristic (ROC) or Precision-Recall Curve can be used. The ROC curve is made based on the value of the confusion matrix, which is to compare the False Positive Rate with the True Positive Rate. To assess and compare the performance of each algorithm, we can look at the area under the curve or AUC (Area Under Curve).

Here are the results of the testing of the 4 Algorima classifications.

Model Ohlson

1. Results of Confusion Matrix and ROC Curve and AUC Long Short Term Memory (LSTM)

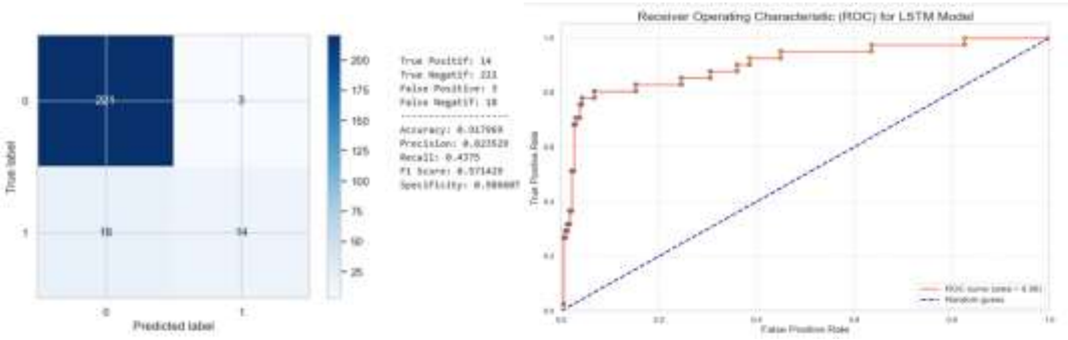

Figure 1 Confusion Metrix and ROC-AUC curve of LSTM Model

Based on the test results presented in Figure 1, the LSTM model shows quite good performance with an accuracy of 91.77%, precision 82.35%, recall 43.75%, F1-score 57.14% and specificity 98.66%. The confusion matrix value shows 14 True Positive, 221 True Negative, 3 False Positive, and 18 False Negative results. The ROC curve with AUC 0.90 indicated excellent discrimination ability. Even though this model shows high accuracy and precision, the relatively low recall value shows that the LSTM model has several weaknesses in detecting all positive cases.

2. Support Vector Machines (SVM) Confusion Matrix and ROC and AUC Curve Results

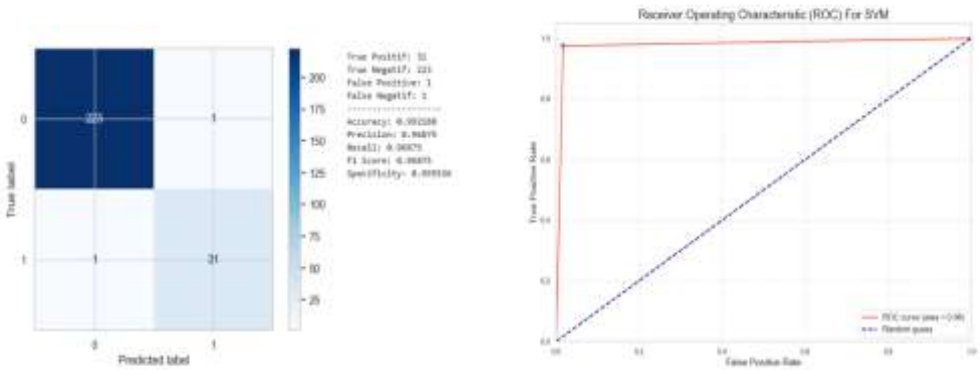

Figure 2 Confusion Metrix and ROC-AUC curve of SVM Model

Based on the test results as presented in Figure 2, the Support Vector Machine (SVM) model shows excellent performance with an accuracy of 99.22%, precision and recall of 96.88% respectively, and an F1 Score of 96.88%. With only 1 error for each False Positive and False Negative, and Specificity 99.55%. It can be concluded that this model is very effective in classifying data. The ROC curve showed an AUC of 0.98, indicating almost perfect discrimination ability. Overall, this model is very reliable in classification with very minimal prediction errors.

3. Results of Confusion Matrix and ROC Curve and AUC Random Forest

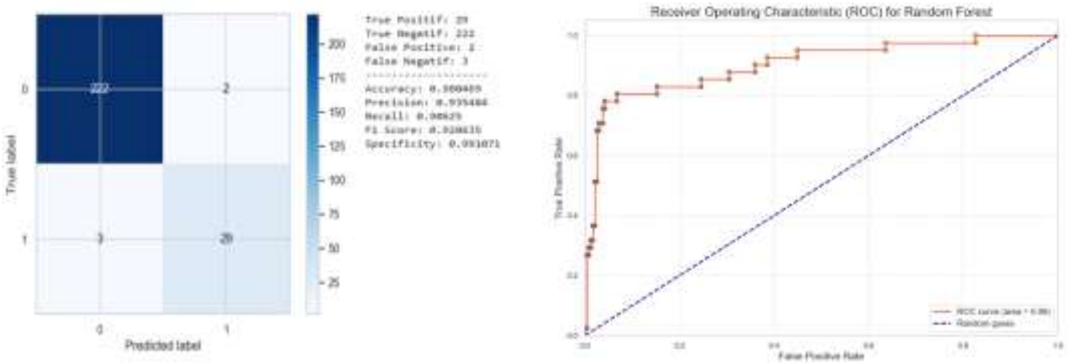

Figure 3 Confusion Metrix and ROC-AUC curve of Random Forest Model

Based on the test results as presented in Figure 3, the Random Forest model shows very good performance with an accuracy of 98.04%, precision of 93.55%, and recall of 90.62%. The confusion matrix value shows 29 True Positive, 222 True Negative, 2 False Positive, and 3 False Negative. F1-Score is 92.06% and specificity reaches 99.11%. The ROC curve with AUC 0.90 indicated excellent discrimination ability.

4. Results of Confusion Matrix and ROC Curve and AUC XGBoost

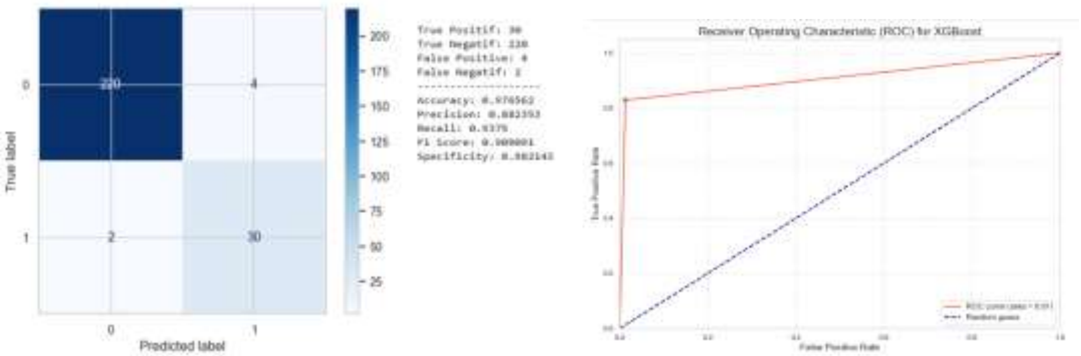

Figure 4 Confusion Metrix and ROC-AUC curve of the XGBoost Model

Based on the testing results presented in Figure 4, the XGBoost model shows good performance with an accuracy of 97.57%, precision of 88.24%, and recall of 93.75%. The confusion matrix value shows 30 True Positive, 220 True Negative, 4 False Positive and 2 False Negative. F1-Score is 90.09% and specificity reaches 98.21%. The ROC curve with an AUC of 0.91 indicates excellent discrimination ability although it performs slightly below the SVM model.

Model Altman

5. Results of Confusion Matrix and LSTM ROC and AUC Curves

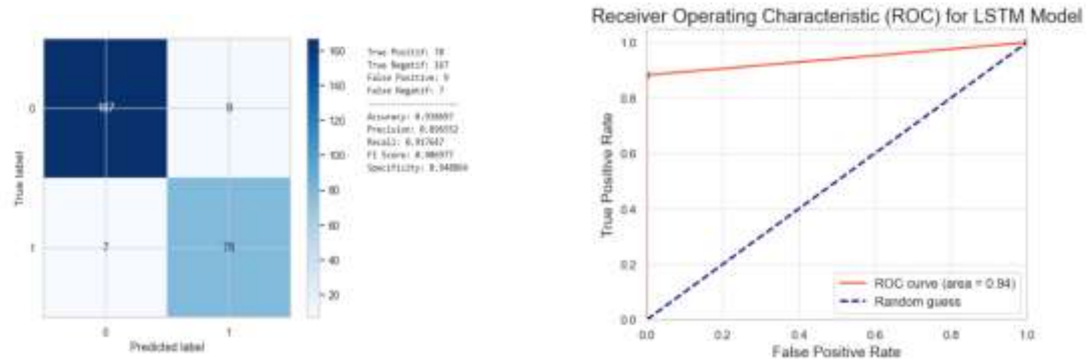

Figure 5 Confusion Metrix and ROC-AUC curve of LSTM Model

Based on the test results in Figure 4.5, the LSTM model shows very good performance with an accuracy of 93.86%, precision 89.65%, and recall 91.76%. The confusion matrix value shows 78 True Positive, 167 True Negative, 9 False Positive, and 7 False Negative. F1-Score is 90.69% and specificity reaches 94.88%. The ROC curve with AUC 0.94 indicated excellent discrimination ability. This model proved to be very reliable in classification with little prediction error which shows that the LSTM model has excellent performance in detecting positive and negative cases.

6. Results of Confusion Matrix and SVM ROC and AUC Curves

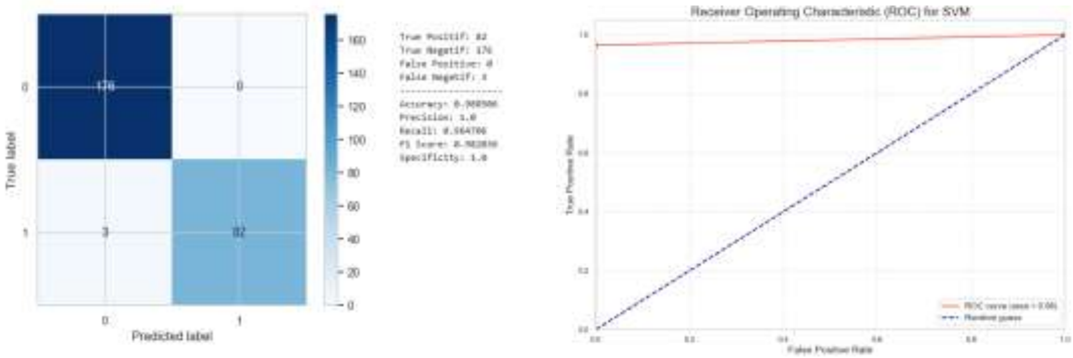

Figure 6 Confusion Metrix and ROC-AUC curve Model SVM

Based on the results of the examination in Figure 6, the SVM model shows very good performance with an accuracy of 98.85%, precision of 100% and recall of 96.47%. The confusion matrix value shows 82 True Positive, 176 True Negative, 0 False Positive, and 3 False Negative. F1-Score is 98.20%, and specificity reaches 100%. The ROC curve with AUC 0.98 indicated excellent discrimination ability. This model proved to be very reliable in classification with little prediction error, which indicates that the SVM model has excellent performance in detecting positive and negative cases.

7. Results of Confusion Matrix and ROC Curve and AUC Random Forest

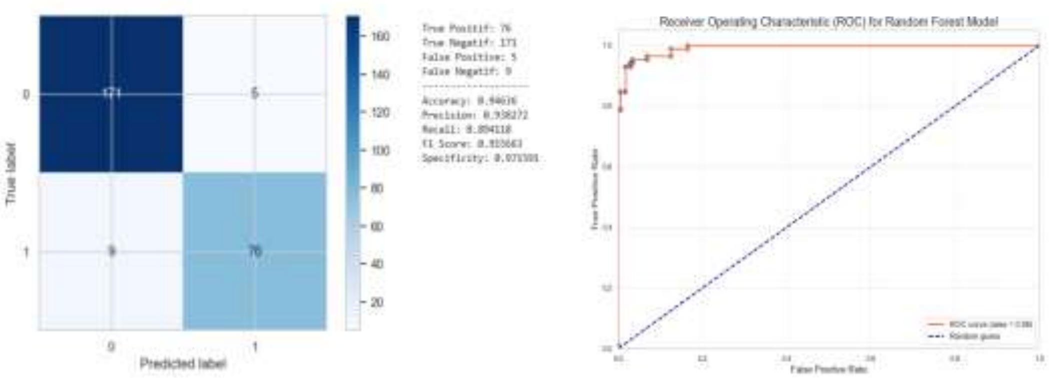

Figure 7 Confusion Metrix and ROC-AUC curve of the Random Forest Model

Based on the test results in Figure 7, the Random Forest model shows very good performance with an accuracy of 94.63%, precision 93.82%, and recall 89.41%. The confusion matrix value shows 76 True Positive, 171 True Negative, 5 False Positive, and 9 False Negative. F1-Score is 91.56% and specificity reaches 97.15%. AUrC Random Forest serbersar 0.99. AUrC = 0.99 means that the True Positive Rater result is always close to 1 compared to the Falser Positive Rater value. This shows that the SVM classifier can very well differentiate between all positive and negative classes correctly. The higher the AUC, the better the model performance in distinguishing positive and negative classes

8. Results of Confusion Matrix and ROC Curve and XGBoost

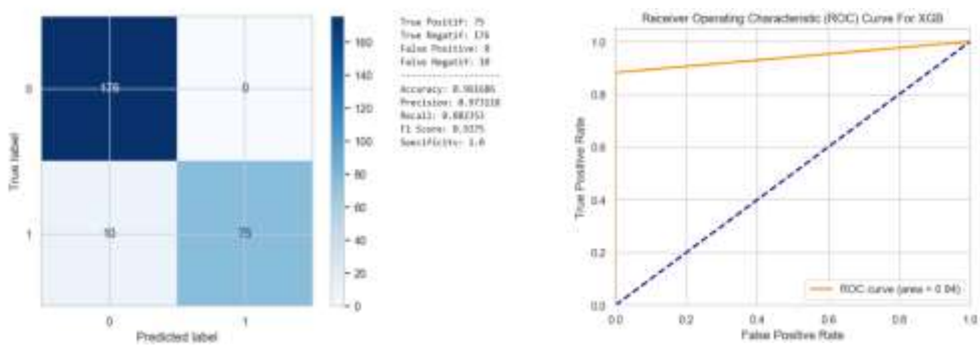

Figure 8 Confusion Metrix and ROC-AUC curve of the XGBoost Model

Based on the testing results as presented in Figure 8, the XGBoost model shows excellent performance with an accuracy of 96.16%, precision of 97.31%, and recall of 88.23%. The confusion matrix value shows 75 True Positive, 176 True Negative, 0 False Positive, and 10 False Negative. F1-Score is 93.75% and specificity reaches 100%. The ROC curve with AUC 0.94 indicated excellent discrimination ability. This model is very reliable in classification tasks with little prediction error showing excellent performance in detecting positive and negative cases with very low error rate.

Table 3 AUC Evaluation Results

Algoritma	AUC	
	Ohlson (%)	Altman (%)

LSTM	0.90	0.94
Support Vector Machine (SVM)	0.98	0.98
Random Forest	0.90	0.99
XGBoost	0.91	0.94

By using 2023 data as new data as many as 147 samples, the accuracy of each model in predicting the Oshlon model and the Altzman Model is obtained as follows:

Table 4
Comparison of Model Prediction Accuracy

Model	Label			Prediction Accuracy							
	Distress	No Distress	Total	SVM	%	XGboost	%	Random Forest	%	LSTM	%
Oshlon	17	130	147	130	88	132	90	130	88	134	91
Altman	44	103	147	112	76	111	75	109	74	113	77

Ohlson Model: Of the 147 companies tested, the SVM and Random Forest machine learning techniques predicted 130 companies correctly (88% accuracy), XGBoost achieved 90%, and LSTM performed best with 91%.

Altman Model: SVM had 76% accuracy, XGBoost 75%, Random Forest 74%, and LSTM best with 77%.

Based on Table, LSTM has the best performance in predicting bankruptcy on 2023 data with an accuracy of 91% for the Ohlson Model and 77% for the Altman Model. This result is different from the 2013-2022 data, where SVM is considered the best. Causes of these differences include differences in sample sizes, overfitting to old data, model complexity and learning capabilities, and changing economic conditions.

Conclusion

Based on the analysis, several main conclusions are as follows. Using data from 2013-2022, Support Vector Machine (SVM) produces the best bankruptcy prediction model based on accuracy, precision, specificity, F1-Score, and recall for the Altman Model and Ohlson Model, demonstrating the effectiveness of SVM in predicting old data. Using new data from 2023, Long Short Term Memory (LSTM) shows the best performance with the highest prediction accuracy of 91% for the Ohlson Model and 77% for the Altman Model, demonstrating the ability of LSTM to handle variations and patterns in new data. Accurate prediction models help stakeholders make better decisions, reduce financial risks and optimize company profits, and create a stable and responsive business environment. This research has several limitations. Since the required GDP (Gross Domestic Product) price index data is not available on the Statistics Agency website, the GDP values for each year are calculated independently using the 2010 GDP values as the base year. The number of research samples is also limited during the 2013-

2023 period. Suggestions for future research include selecting variables that are more relevant and informative in predicting corporate bankruptcy, as well as adding a longer annual deficiency reporting period for a more in-depth and accurate analysis.

Bibliography

Almas, Nadiva Jihan. (2023). *Dampak Industri Manufaktur Terhadap Pertumbuhan Ekonomi: Komparatif Korea Selatan Dan Indonesia*. Universitas Islam Indonesia.

Ariyanto, Anto. (2017). *Literature Review: Agroindustri Kelapa Sawit, Dampaknya Terhadap Ekonomi dan Daya Saing Indonesia*.

Creswell, Antonia, White, Tom, Dumoulin, Vincent, Arulkumaran, Kai, Sengupta, Biswa, & Bharath, Anil A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, *35*(1), 53–65.

Joshi, Shreya, Ramesh, Rachana, & Tahsildar, Shagufta. (2018). A bankruptcy prediction model using random forest. *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 1–6. IEEE.

Kothuru, Vivek, Jha, Raman Kumar, Ranjit, Sasanapuri, Reddy, Bommireddy Vijay Kumar, Roy, Sagnik, & Sudheer, Sachin. (2022). Prediction of Bankruptcy of a company using machine learning techniques. *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 797–800. IEEE.

Madjid, Verina Ardiyanti, Mahdi, Fauzan Akmal, Lukito, Cahyo Adi, Nofri, Dinda Parwita Aulia, & Prasvita, Desta Sandya. (2021). Pengaruh Principal Component Analysis terhadap Akurasi Model Machine Learning dengan Algoritma Artificial Neural Network untuk Prediksi Kebangkrutan Perusahaan. *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer Dan Aplikasinya*, *2*(2), 110–119.

Muta'ali, Luthfi. (2019). *Dinamika peran sektor pertanian dalam pembangunan wilayah di Indonesia*. UGM PRESS.

Shetty, Vijith Vittal, & Kellarai, Adithi. (2022). Comprehensive review of hepatocellular carcinoma in India: current challenges and future directions. *JCO Global Oncology*, *8*, e2200118.

Siswoyo, Bambang. (2020). MultiClass decision forest machine learning artificial intelligence. *Journal of Applied Informatics and Computing*, *4*(1), 1–7.

Sulastri, Desra Afri. (2014). Pengaruh Volatilitas Arus Kas, Volatilitas Penjualan, Besaran Akrual Dan Tingkat Hutang Terhadap Persistensi Laba (Studi Empiris pada Perusahaan Manufaktur yang Terdaftar di BEI Tahun 2009-2012). *Jurnal Akuntansi*, *2*(3).

Swari, Areta Betari, & Pristiana, Ulfi. (2020). Pengaruh Makro Ekonomi Terhadap Nilai Perusahaan Dengan Kinerja Keuangan Sebagai Variabel Intervening Pada Perusahaan Sub Sektor Kontruksi dan Bangunan yang Terdaftar di BEI Tahun 2016-2018. *Jurnal Ekonomi Manajemen (JEM17)*, *5*(2), 107–127.

Wibowo, Andi. (2012). *Analisis pengaruh produk domestik bruto (pdb) dan investasi riil* *sektor industri terhadap penyerapan tenaga kerja pada industri pengolahan di indonesia tahun 1985–2011*.