Comparison of Financial Distress Prediction Model Accuracy Using Support Vector Machines and Discriminant Analysis Methods

Herlina*, Siti Mundari

Department of Industrial Engineering, Universitas 17 Agustus 1945 Surabaya, Indonesia
***Corresponding author**: herlina@untag-sby.ac.id

Abstract

Financial distress is a stage before the company goes bankrupt. For this reason, the ability to predict financial can be useful information for companies and investors. This information is useful for companies to be able to improve their financial condition so that the company does not go bankrupt. For investors, this information is useful to avoid investor losses in capital investment. Studies on financial distress have been conducted for a long time, starting with using statistics until now being developed using artificial intelligence methods. The purpose of this study was to compare the accuracy of the financial distress prediction model for publicly traded manufacturing companies in miscellaneous industry sectors listed on the Indonesia Stock Exchange using the data mining method, namely Support Vector Machines, which is one of artificial intelligence method and the statistical method, namely Discriminant Analysis. From the research results, the two methods provide equally good accuracy. Based on the processed data, the accuracy of the two methods is 100%.

Keywords: *support vector machines, discriminant analysis, data mining, financial distress, miscellaneous industry*

Abstrak

Financial distress atau kesulitan keuangan merupakan sebuah tahapan sebelum perusahaan mengalami kebangkrutan. Oleh karena itu, kemampuan dalam memprediksi *Financial Distress* dapat menjadi sebuah informasi yang berguna bagi perusahaan dan investor. Informasi ini berguna bagi perusahaan untuk dapat memperbaiki kondisi keuangannya sehingga perusahaan tidak sampai mengalami kebangkrutan. Bagi investor, informasi ini berguna untuk menghindari kerugian investor dalam penanaman modal. Penelitian mengenai financial distress sudah dilakukan sejak lama, mulai dengan menggunakan metode- metode statistik sampai saat ini terus dikembangkan dengan menggunakan metode-metode kecerdasan buatan. Tujuan dari penelitian ini adalah membandingkan keakuratan model prediksi *financial distress* pada perusahaan manufaktur terbuka pada sektor aneka industri yang terdaftar di Bursa Efek Indonesia dengan menggunakan metode data mining yaitu Support Vector Machines yang merupakan salah satu metode kecerdasan buatan dan metode statistik yaitu Analisis Diskriminan. Dari hasil penelitian, kedua metode tersebut memberikan akurasi yang sama baiknya. Berdasarkan data yang diolah, akurasi kedua metode tersebut adalah 100%.

Kata Kunci: *support vector machines, analisis diskriminan, data mining, financial distress, aneka industri.*

1. Introduction

Financial distress can be experienced by all companies and the opportunity is greater when there is an economic crisis in the country where the company operates [5]. Financial distress conditions that can be predicted early on can provide benefits for the company because the company can try to improve its condition both in terms of marketing and production by implementing special strategies [4]. Prediction of financial distress also provides useful information for investors and creditors as material for consideration in decision making [8]. Because of the benefits for companies and various interested parties, research on the prediction of financial distress is an interesting research topic and continues to be developed.

The research began with a statistical method which was first carried out by Altman in 1968 using the Multiple Discriminant Analysis (MDA) method [1]. Another statistical method used is Logistic Regression. From studies using statistical methods [15] where one of the drawbacks is that there are still assumptions about normality and linearity, then studies are being developed using data mining and artificial intelligence techniques, starting from Decision Trees, Artificial Neural Networks, Support Vector Machines. One of the advantages of data mining techniques is that there are no assumptions about normality and linearity as found in statistical methods [9]. Data mining techniques use classification techniques that will classify data

into their respective classes, in this case the class of companies experiencing financial distress and the class of companies that are not experiencing financial distress.

In previous research [2] suggested that Support Vector Machines (SVM) is a classification method that is quite reliable for classification cases. The use of the SVM method for predicting financial distress that has been studied by [2] was applied to companies in South Korea with good results. Likewise applied to companies in China and also got good results.

SVM is an algorithm proposed by Vapnik in 1995. SVM is classified as a new classification method and has been widely used as a method in a number of studies, such as pattern recognition, regression, and estimation. According to [11], SVM is in the same class as ANN (Artificial Neural Network) but in many implementations it is proven that SVM provides better results than ANN in terms of the solutions achieved. ANN finds a solution that is locally optimal while SVM finds a solution that is globally optimal. The SVM method will map the input data into a high-dimensional space which will later obtain a linear separator function that will separate the data into different classes. To determine which of the infinitely possible separator functions (classifier/hyperplane) is optimal for separating two types of objects, SVM employs the input from the training data. Maximizing the margin or distance between two groups of items of different classes yields the optimal hyperplane. SVM is applicable to both linear and non linear data. For classification cases, where the data is not linear, you can use the Kernel method.

This study will compare the accuracy of the financial distress prediction model for public listed manufacturing companies on the Indonesia Stock Exchange specifically for miscellaneous industry sectors using the Support Vector Machines (SVM) and Discriminant Analysis methods[7].

2. Material and Methods

This study used a sample of publicly traded manufacturing companies in miscellaneous industry listed on the Indonesia Stock Exchange (IDX) [3]. Companies that are the object of observation in this study are companies that have complete financial statements. The financial reports that will be used in this study are annual financial reports for the period of three years. Data on the financial reports of observed companies can be obtained through the Indonesia Capital Market Directory (ICMD). Companies in financial distress should be labelled -1, and those not in financial distress should be labelled 1. If a company's financial accounts show that it satisfies one or more of the following conditions, it is said to be in financial trouble [6]:

1. Working capital that is negative
2. Operating profit that is negative
3. Net profit that is negative

The next step is the selection of variables as indicators of financial distress. Before the values of these variables are entered into the model, it is necessary to carry out data transformation or what is known as data preprocessing. This data transformation needs to be done to make the data range that was previously not the same for each variable to be uniform according to a predetermined range. The technique that will be used in transforming the data in this study is scaling. Scaling is the procedure of changing data so that it is on a certain scale. In this study a scale (-1.1) will be used, in which case the lower limit is -1 and the upper limit is 1.

The next step is to determine the Training Data and Testing Data which will be randomly selected datasets as training data and testing data. The ratio between training data and testing data to be used is 2:1. The next step is to create a financial distress prediction model using the Support Vector Machine (SVM) and Discriminant Analysis methods. Comparing the results of SVM and Discriminant Analysis in terms of accuracy.

3. Results and Discussion

There are 14 publicly listed manufacturing companies in Indonesia in miscellaneous industry sectors that have complete financial reports in the observation period. The number of companies experiencing financial distress (negative class) is 5 companies while the number of companies that are not experiencing financial distress (positive class) is 9 companies.

The variables [10] that will be used in this study include financial variables with several financial ratios [12], including liquidity ratios, solvency ratios, profitability ratios and investor ratios [14]. The list of variables and their formulas can be seen in Table 1.

Table 1. List of variables

Criteria	Sub Criteria	Description
liquidity ratios	Working Capital – to - Total Asset	(Current Assets – Current Liabilities) : Total Assets
	Current Ratio	Current Assets : Current Liabilities
solvency ratios	Debt-to-Equity	Total Liabilities : Total Equity
	Leverage Ratio	Total Liabilities : Total Assets
profitability ratios	Gross Profit Margin	Gross Profit : Sales
	Net Profit Margin	Net Income : Sales
	Operating Profit Margin	Operating income : Sales
	ROE	Net Income : Total Equity
	ROA	Net Income : Total Assets
investor ratios	EPS	(Net Profit – preferential stock dividends) : the number of ordinary shares outstanding
	Book Value per Share	(Total Stockholders' Equity – Preferred Stock) : the number of ordinary shares outstanding

The company's liquidity [13] describes the company's ability to meet its short-term obligations and also to measure the efficiency of using current assets. The solvency ratio is used to measure a company's ability to meet its long-term obligations and also to measure the level of protection provided by the company to its long-term creditors. Investor ratios are certain profitability ratios that are of concern to investors.

Working Capital-to-total assets, is a comparison between the difference between total current assets and current liabilities with total assets. This ratio is used to measure the liquidity of the company. Current Ratio, is a comparison between current assets and current liabilities. This ratio is used to measure the extent to which a company is able to pay its short term debts.

Debt-to-equity, is a comparison between total debt (debt) with total assets (assets). This ratio provides an overview of the company's capital structure to see the level of risk of uncollectible debt. Leverage ratio, is a comparison between total debt (liabilities) and total assets (assets). This ratio is used to measure how much total assets are financed by total existing debt. The higher this ratio, the greater the risk faced by the company.

Gross profit margin, is a comparison between gross profit and sales. This ratio measures production efficiency and determines the selling price. Net profit margin, is a comparison between net profit and sales. This ratio is used to measure how much the proportion of net profit received by a company comes from sales. Operating Profit margin, is a comparison between operating income and sales. This ratio is one of the profitability ratios used to measure a company's ability to generate income to maintain its business continuity. Return On Equity (ROE), is a comparison between net income and total equity. This ratio measures the company's ability to generate profits that will benefit shareholders. Return On Assets (ROA), is a comparison between net income and total assets. This ratio measures a company's ability to utilize its assets to earn profits.

Earning per Share (EPS), is a comparison between the difference between net profit and dividends from special shares and the number of outstanding shares. This ratio is used to measure the amount of profit that is entitled to each holder of one share of common stock. Book value per share (BV), is a comparison between the difference between shareholder equity and preferred stock and the number of ordinary shares outstanding. This ratio shows the number of stockholders' equity associated with each outstanding share.

The financial distress prediction model used in this study, namely Support Vector Machines uses data mining techniques, where data mining or also known as knowledge discovery in databases is an activity that includes the collection and use of historical data to find regularities, patterns or relationships in large data sets. [7] Based on the training data used in the model, patterns will be found to predict label values from data testing. So to get predictive results whether a company is experiencing financial distress or not in this research model, it is not necessary to have special parameter values from financial ratios used as indicators of financial distress. Through a number of training data used, when new data is entered to be tested, the predictive label of the test data will be obtained by reading patterns from a number of training data that have been used to train the model.

In this study, the comparison between testing data and training data to be used is 2:1. According to [5], a balanced dataset will provide a greater degree of accuracy than a dataset that is unbalanced between positive and negative classes. Because the dataset between the positive class and the negative class is not balanced, the training data and testing data will be balanced by taking some of the positive class data randomly as much as the number of negative class datasets.

The training data used in this study were 4 companies from the positive class and 4 companies from the negative class. The testing data used in this study were 2 companies from the positive class and 2 companies from the negative class. The predetermined training data and testing data will be tested computationally with MATLAB software using the SVM method. The results of model testing with several types of Kernels can be seen in **Table 2**.

Table 2. The results of testing the dataset using SVM

Kernel	Misclassification (%)
Linear, with parameter C=100	0
Polynomial with parameter degree 4 and C=100	0
RBF with parameter $\sigma = 2$, C=100	0

From the test results with the Support Vector Machines (SVM) method, a misclassification result of 0% was obtained. This means that the prediction accuracy rate is 100%. From the results of data processing using SPSS software for Discriminant Analysis, the results obtained can be seen in **Fig. 1**.

Canonical Discriminant Function Coefficients

	Function 1
Working Capital to Total Assets	7.428
(Constant)	-1.327

Unstandardized coefficients

Fig. 1: SPSS Output Canonical Discriminant Function Coefficients

Model testing using Discriminant Analysis yields the discriminant function equation: D = -1.327 + 7.428 Working Capital to Total Assets.

Classification Results[a,c]

		Kelas	Predicted Group Membership		Total
			Financial Distress	Non Financial Distress	
Original	Count	Financial Distress	5	0	5
		Non Financial Distress	0	9	9
	%	Financial Distress	100.0	.0	100.0
		Non Financial Distress	.0	100.0	100.0
Cross-validated[b]	Count	Financial Distress	5	0	5
		Non Financial Distress	1	8	9
	%	Financial Distress	100.0	.0	100.0
		Non Financial Distress	11.1	88.9	100.0

a. 100.0% of original grouped cases correctly classified.

b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

c. 92.9% of cross-validated grouped cases correctly classified.

Fig. 2: SPSS Output Classification Results

Based on the information from the Classification Results **Fig. 2**., it provides information that the prediction accuracy rate is 100%. This means that the resulting discriminant equation is suitable for use in predicting whether a company enters the financial distress or non-financial distress group based on information on the value of Working Capital to Total Assets.

4. Conclusion

From the results of data processing using MATLAB software for the Support Vector Machines (SVM) method, it was found that all testing data could be categorized correctly so as to achieve 100% accuracy. Likewise, from the results of data processing using SPSS software for the Discriminant Analysis method, the results showed that 100% of the data was classified correctly. So in this research the results obtained were that the comparison of accuracy in Financial Distress predictions using the Support Vector Machines and Discriminant Analysis methods both gave the same good results with an accuracy of 100% for each method.

5. Acknowledgment

6. References

[1] Arif, M.F., 2022, Analisis Perbandingan Model Pendeteksi Financial Distress, Jurnal Riset Terapan Akuntansi, Vol. 6 No. 1

[2] Chen, B. –T. dan Chen, M. –Y., 2010, Applying Particles Swarm Optimization For Support Vector Machines On Predicting Company Financial Crisis,*International Conference on Business and Economics Research*, Vol. 1, Kuala Lumpur, Malaysia.

[3] Ismiati, E., Wibowo, E., and Utami, S. S., 2020, Pengaruh Rasio Keuangan Terhadap Prediksi Financial Distress (Studi pada perusahaan manufaktur sector aneka industry yang terdaftar di BEI periode 2015 – 2018), Jurnal Ekonomi dan Kewirausahaan, Vol. 20 No. 2

[4] Irfan, F.M. and Supriyanto, R., 2024, Comparative Analysis of Machine Learning and Deep Learning Models Integrated with Altman Z-Score for Financial Distress Prediction in Companies Listed on the Indonesia Stock Exchange (IDX), Ekombis Review: Jurnal Ilmiah Ekonomi dan Bisnis, Vol. 12 No. 2 April 2024 page: 2259 – 2278

[5] Kholifah, N., Djumali, Hartono, S., 2020, Mengukur Financial Distress Dengan Metode Grover, Altman Z-Score, Springate Dan Zmijewski Pada PT. Solusi Bangun Indonesia Tbk, Edunomika, Vol. 4 No.2.

[6] Kurniadi, A., 2021, Analisis Rasio Keuangan Untuk Memprediksi Financial Distress Perusahaan Manufaktur Di BEI, JIMKES Jurnal Ilmiah Manajemen Kesatuan Vol. 9 No. 3

[7] Kusuma, J. and Hadiprajitno, B., 2021, Prediksi Financial Distress Perusahaan Di Indonesia Menggunakan Rasio Keuangan Dan Analisis Diskriminan, DIPONEGORO JOURNAL OF ACCOUNTING, Volume 10, Nomor 4

[8] Nisa, U. Z., 2013, Model Prediksi Finansial Distress Pada Perusahaan Manufaktur Go Public di Indonesia. Thesis Program Magister Bidang Optimasi Sistem Industri Jurusan Teknik Industri Institut Teknologi Sepuluh Nopember.

[9] Pelitawati, D. and Kusumawardana, R.A., 2020, Analisis Komparasi Model Altman, Zmijewski Dan Springate Untuk Memprediksi Financial Distress Pada Perusahaan Yang Terdaftar Di Bursa Efek Indonesia, Jurnal Economics and Sustainable Development Vol. 5 No.02

[10] Riyanti, S. D., 2020, Analisis Rasio Keuangan terhadap Kondisi Financial Distress pada Perusahaan Manufaktur, Jurnal READ, Vol.1 No.2, hal. 56-65.

[11] Santosa, B. dan Willy, P., 2011, *Metoda Metaheuristik Konsep dan Implementasi*, Guna Widya, Surabaya.

[12] Sari, M. and Diana, H., 2020, Financial Ratio Analysis To Predict The Financial Distress Conditions Of Pulp And Paper Companies Listed On The Iindonesia Stock Exchange In 2012-2017 With The Altman Z-Score Model, Research In Accounting Journal, Vol 1(1) 2020 : 32-48

[13] Stephanie at al, 2020, Pengaruh Likuiditas, Leverage Dan Ukuran Perusahaan Terhadap Financial Distress Pada Perusahaan Properti Dan Perumahan, COSTING:Journal of Economic, Business and Accounting, Volume 3 Nomor 2

[14] Suhardoyo and Muktiningrum, 2024, Prediksi Financial Distress Pada Perusahaan Tektil Yang TerdaftarDi Bursa Efek Indonesia (BEI) Melalui Analisis Rasio Keuangan, J-CEKI : Jurnal Cendekia Ilmiah, Vol.3, No.6

[15] Suhendi, A., 2021, Analisis Altman Z-Core pada Perusahaan Manufaktur Sub Sektor Otomotif yang Terdaftar di Bursa Efek Indonesia, Jurnal Gentiaras Manajemen dan Akuntasi, Vol. 13 No. 2.