Comparing the Accuracy of Multiple Discriminant Analyisis, Logistic Regression, and Neural Network to estimate pay and not to pay Dividend

Triasesiarta Nur
Universitas Bina Nusantara

* *triasesiarta.nur@binus.ac.id*

Abstrak

This study compares the accuracy of prediction to estimate the companies dividend policy; in this case, the company will pay or not pay dividends. The models used in this research are Multiple Discriminant Analysis, Logistic Regression, and Neural Network. The samples are divided into two groups, namely companies that always pay and not pay dividends during the 2015-2018 research period, resulting in 256 samples not paying dividends and 128 samples paying dividends. The results showed that the average Neural Network accuracy performance exceeded the other two models. The best predictor of the company's Dividend Policy in this study is Price to Book Value, Stock Price, Firm Cycle, current ratio, ROA and Exchange Rate.

Keywords: Multiple Discriminant Analysis, Logistic Regression, Neural Network, Dividend Policy

Introduction

Dividend policy is a decision whether profits derived by a company will be distributed to shareholders as dividends or will be retained in the form of retained earnings to finance investment in the future. If the company chooses to distribute profits as dividends, it will reduce retained earnings and subsequently reduce total internal resources. Various theories were developed to explain the behavior of corporate dividend policies. These theories include A bird in the hand (Gordon, 1959); (Lintner J., 1956), Agency cost (Jensen & Meckling, 1976), Signaling (Bhattacharya, 1979); Behavioral explanation (H. M. Shefrin & Thaler, 1988)(H. Shefrin & Statmant, 1985); Life Cycle (Grullon, Michaely, & Swaminathan, 2002) and Taxation (Miller & Scholes, 1978). Various studies have also been conducted to analyze the factors that influence dividend policy. These factors include profitability, liquidity, efficiency, leverage, life cycle, ownership and size of the company (Lestari, 2018); (Franc-Dąbrowska, Mądra-Sawicka, & Ulrichs, 2020). The need to be able to predict whether a company will pay dividends or not is essential for investors to determine their investment decision. Companies that pay dividends are considered as companies that have good prospects in the future.

The method used to estimate dividend policy is divided into several types, namely multiple regression, Tobit regression, Logistic Regression, Multiple Discriminant Analysis and more recently through the Neural Network approach. Within the dividend policy framework, size forms the basis of the method to be used. In this matter, the size of the dividend policy are represented by dividend yields, dividends per share; or policies to pay or not pay dividends. In this context, Multiple regression models and Tobit regression models are used when the dependent variables are represented dividend yield and dividend per share. Where as in the case of pay or non-payment of dividends, the Logistic Regression model

Triasesiarta Nur

(LR), and the Multiple Disriminant Analysis model (MDA). Various studies indicate specifically that sample selection will exclude companies that do not pay dividends. This condition affects the essence of dividend policy, where the policy of not paying dividends is a decision taken by the company.   The reason in using the approach to eliminate companies that do not pay dividends from the sample, is to avoid the data distribution skewness. As statistical tool Tobit regression uses censored data namely because the use of ordinary least square in a skew data distribution will cause bias in result (Tobin, 1958); (Greene, 2000).

Neural Network (NN) is a model that follows the workings of the human brain, consisting of interconnected neurons that are used to provide solutions to various problems in various fields of science such as statistics, technology, and economics. Some studies show neural networks have more accurate predictions than other models (Abdou, Pointon, El-Masry, Olugbode, & Lister, 2012); (Laoh, 2019). This study aims to compare three models, i.e. Multiple Discriminant Analysis Model (MDA), Logistic Regression Model (LR), and Neural Network (NN), for predicting corporate dividend policy.

Theoritical Framework
Multiple Discriminant Analysis

Multiple discriminant analysis (MDA) is part of a multivariate statistical analysis that aims to separate several groups of data that have been grouped by forming discriminant functions. This statistical technique is used in dependency relationships between response variables and explanatory variables. Discriminant analysis is used in cases where the response variable is qualitative data, and the explanatory variable is quantitative data (Coulombe, 1985). MDA is appropriately used to examine the company's dividend policy, which is to comprehend the distinguishing factors between companies that pay and do not pay dividends (Ismiyanti, 2005). The model of discriminant analysis is as follows:.

$$Z = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_i X_i \qquad (1)$$

where, Z= discriminant score, α = a constant term, βi = the discriminant coefficient or weight of the variable, Xi = predictor or independent variable, i = number of predictor variables; i = 1,2,3,....k.

Logistic Regression

Logistic regression (LR) is one type of regression that connects one or several independent variables with the dependent variable in the form of a category; 0 and 1. The types of independent variables in this category distinguish logistic regression from multiple regression or other linear regression. LR models the probability of an outcome based on individual characteristics. The model of LR is as follows:

$$log \left(\frac{\pi}{1-\pi} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ...\beta_m x_m \qquad (2)$$

π is the probability of an event (e.g. paying dividends) and β1 is the regression coefficient related to the reference group while x1 is an explanatory variable (Sperandei, 2014).

Machine Learning

An approach in Artificial Intelligence that is widely used to replace or simulate human behavior to solve problems or automate. As the name implies Machine Learning (ML) that tries to mimic how human or intelligent beings learn and generalize. There are at least 2 main applications in ML, namely classification and prediction. The hallmark of ML is the process of training, learning and training. Therefore, ML requires data to be studied which is called training data. Classification is a method in ML that is used by machines to sort or classify objects based on certain characteristics as humans try to distinguish objects from one another. While predictions or regressions are used by machines to guess the output of an input based on data that has been learned in training. The most popular ML methods are decision making systems, support vector machines and neural networks.

Neural Network.

Artificial neural network systems (NN) are computer algorithms that can be 'trained' to mimic neural networks in the human brain (Dorsey, Edminister, & Johnson, 1995). The network consists of a large number of basic processing units that are interconnected to process information. The results of network processing come from the collective behavior of the units and depend on how the units interact with each other (Altman, Marco, & Varetto, 1994). By processing and evaluating interactions in previously complex data sets, the neural network tries to assign the right weights to each input to enable the correct reduction of the final result. These input weights are aided by the 'genetic algorithm' optimization procedure, which simulates the predictive power of the model under a large number of scenarios and allows the best weighting scheme to survive and reproduce from one generation to the next.

NN is one of the techniques in machine learning that mimics the human nerve, which is a fundamental part of the working brain. NN consists of the input and output layers. Each layer consists of several neurons that have an activation function that determines the output of the unit. Hidden layers can be added to increase the capabilities of the NN. NN is trained using the training data. The more training data, the better the performance of the NN. The ability of NN is also limited to the number of layers, the more the number of layers, the higher the capacity of the NN.

Methodology

This study uses secondary data taken from financial statements of public companies from the manufacturing sector listed on the Indonesia Stock Exchange in the period 2015-2018. Based on the respective dividend policy two groups of companies are formed i.e companies that continue to pay dividends, and alternately companies that never pay dividends.

Research Model

This research tests the prediction accuracy of factors that determine the company's policy to pay or not pay dividends. Three models are used: Multiple Discriminant Analysis, Logistic Regression and Neural Network. As predictors 24 financial ratios consisting of liquidity, effectiveness, profitability, leverage, market ratios are used, in addition to ratios these factors relating to size, age, company life cycle and macroeconomic conditions are included. The selection of the best predictor is performed by using stepwise discriminant

analysis in MDA (Hertz et al, 1991). In the process, all predictors are entered into the model and then removed gradually if not statistically significantly influences discrimination. In LR model, the selection of the best predictor is performed using the forward (conditional) method. Whereas on neural networks the choice of predictors is based on the MDA model and LR model.

Comparison of Prediction Accuracy

Comparison of the accuracy of predictions the company hat pay or not pay dividends from the three models using three metrics namely: accuracy, precision, and sensitivity (Devi & Radhika, 2018). There is a fourth measure, namely specificity. However, as the sample size of the two groups of companies is not the same renders the specificity results not being relevant in this research. The calculations of those metrics are based on the classification of comparison between the predicted results and observations data for each model. The outcome of that classification are groups of samples namely: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) calculations. The formula for the three metrics is as follows:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$precision = \frac{TP}{TP+FP} \tag{4}$$

$$sensitivity = \frac{TP}{TP+FN} \tag{5}$$

Result and Discussion
Statistics Descriptive

Data in table 1 shows descriptive statistics of 2 groups of companies that pay and do not pay dividends. The data shows that companies that pay dividends tend to be more liquid, profitable, have higher prices, and are in a cycle where retained earnings against more assets. Even so, the company that pays dividends are companies with a lower-Price to Book Value, and occurs when the Rupiah exchange rate against the US Dollar tends to weaken.

Table 1. Statistics Descriptive

Ratios	Variable	Paying Dividends		Not Paying Dividend	
		Mean	SD	Mean	SD
Liquidity	CR	263.5239	189.1805	166.4598	133.4619
Profitability	ROA	8.0179	7.5715	-.1989	8.7824
Price	lnPrice	7.2941	1.3757	5.8334	1.5159
Firm Value	PBV	3.2075	6.2687	5.1123	41.3618
Firm Cycle	LnKurs	.3529	.2036	-.1510	.7510
Macroeconomy		9.5263	.0306	9.5439	.0464

Source: Search Results. (Multiple Discriminant Analysis Results)

Triasesiarta Nur

MDA test results show a significant difference in the group of companies that pay dividends and those that do not pay dividends, indicated by the output value of Wilks Lambda, which is below 0.05. The best five predictors are obtained by the stepwise method. Those predictors are natural logarithm of price Price to Book Value, Current Ratio, Return On Assets, Exchange Rate, and Retain Earning to Total Assets. The canonical value is 0.692, which means 77.51% of the variance of the independent variables can explain the discriminant model that is formed. The following is a dividend policy model based on MDA:

$$Z = 8{,}153 + 0.325 lnprice - 0.08PBV + 0.02CR + 0.43\ ROA - 0.01ER + 0.323RETA \qquad (6)$$

with a cutting score of -0.513 for companies that pay dividends and 1,047 for companies that do not pay dividends. Table 2 shows the classification results of the percentage of cases predicted accurately

Table 2. Observed-Predicted Dividend Policy MDA Classification

Dividend Policy	Predicted		Total
	Not Paying	Paying	
Not Paying	227	29	256
Paying	42	79	128
			384

Source: Search Result (Logistic Regression Results)

The logistic regression model is used to classify the company's dividend policy, whether the company will pay or not pay dividends. Logistic regression analysis examines the relationship between the dependent variable (dividend policy) and a group of independent variables that will provide predictive results regarding the probability of companies paying dividends. Forward Conditional method was performed to obtain the best predictors as follows in equation 7 :

$$\frac{LnP}{1-P} = 13.021 + 0.479 lnprice - 0.03PBV + 0.154ROA - 0.01Kurs + 2.384RETA \qquad (7)$$

The value of Nagelkerke R Square is 0.641, which means 64,1 % of the variation of the dividend policy is explained by model. The value of Hosmer and Lemeshow test is 0.10 which means the model is fit. The classification of predictive and observed value of LR model can be seen in Table 3.

Table 3. Observed-Predicted Dividend Policy LR Classification

Observed		Predicted		Total
		Dividend Policy		
		Not Paying	Paying	
Dividend Policy	Not Paying	228	28	256
	Paying	50	78	128
				384

Source: Search Results (Neural Network Results*)*

Triasesiarta Nur

The dividend policy was also analyzed by applying the Neural Network. In this model, the research samples are divided into 302 training samples and 78 testing samples. The network has five inputs where each input node represents one variable, one hidden layer with three neurons. This model obtains an incorrect prediction for the training sample of 14.2% and 11.5% for the testing sample. Based on the neural network, the best predictors are ROA, PBV, RETA, Price, and Exchange Rates, respectively. The following table is the result of a predictive classification of the NN:

Tabel 4. Klasifikasi Observed-Predicted Dividend Policy Neural Network Classification

Jenis Sample	Observed Dividend Police	Predicted Dividend Policy		Total
		Not Paying	Paying	
Training	Not Paying	167	33	199
	Paying	10	92	102
Testing	Not Paying	50	7	57
	Paying	6	20	26

Source: Search Results

Neural Network prediction accuracy performance exceeded both MDA and LR for all three measurements as shown in figures 1-4.

Comparison of Accuracy Performance

Accuracy represents comparison between the percentage of cases that are classified correctly. Accuracy is usually used as an estimate of classification performance. NN model has an accuracy performance result above the other two models. For training samples, the value of accuracy level is 85.76% while for tests samples it reaches 88.46%. As a comparison, the MDA model has an accuracy rate value of 79.69% while the LR model has a slightly better accuracy level value of 80.79%. The next metric is precision which is the percentage of cases that are classified positively, as seen in figure 1.
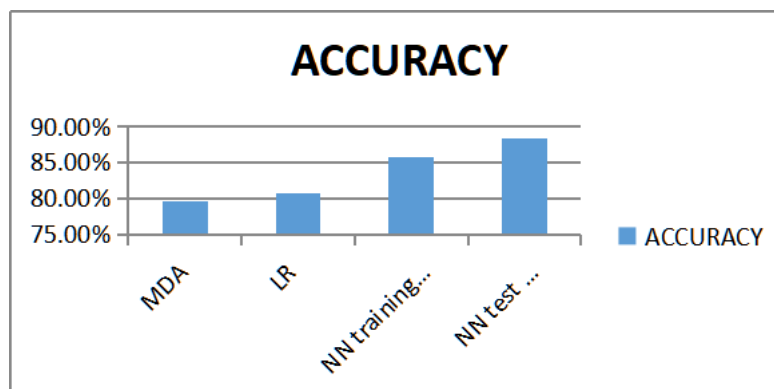

Figure 1. *Comparison of Accuracy Performance*

Comparison of Precision Performance

In the precision metric, the Neural Network model's prediction accuracy performance marginally exceeds that of MDA and LR, the value of precision obtained from training

Triasesiarta Nur

samples is 73.81% and 74.07% test samples, while comparison precision value obtained from MDA models is 73.58%, and 73.64% for LR model. The next metric is sensitivity, defined as the ratio of the number of true negative cases to the total number of negative cases [23], the comparison of sensitivity of the three models can be seen in Figure 2 as follows,
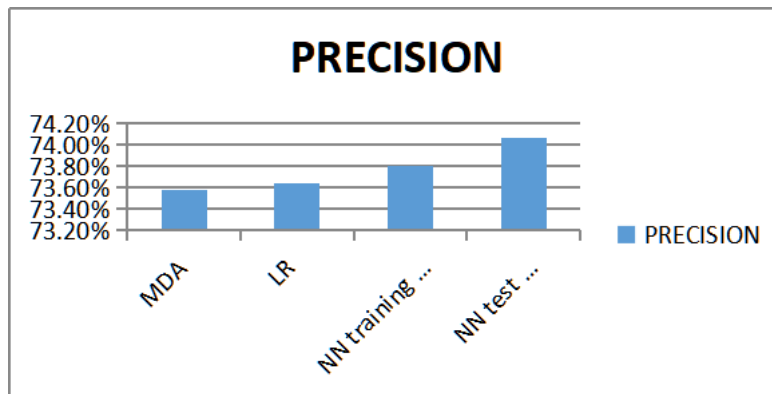

Figure 2. *Comparison of Precision Performance*

Comparison of Sensitivity Performance

Similar to the previous two metrics, the sensitivity performance of neural networks it also exceeds MDA and LR models. Both Neural Network samples have a sensitivity performance of 90.29% for training sample and 90.91% for the test sample, significantly greater than the MDA performance of 60.94% and the LR value of 64.80% respectfuly.
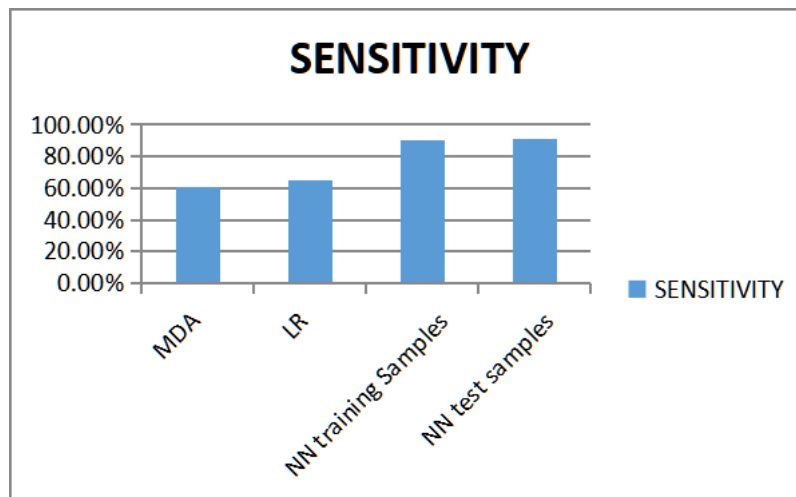

Figure 3. *Comparison of Sensitivity Performance*

Overall average prediction

Performance of the three metrics i.e. accuracy, precision and sensitivity can be seen in Figure 4. Neural Network has the average value of predictive accuracy performance are

Triasesiarta Nur

83.29% for sample training and 84.48% for testing sample. The average predictive performance for MDA is 71.48% and 73.22 for LR
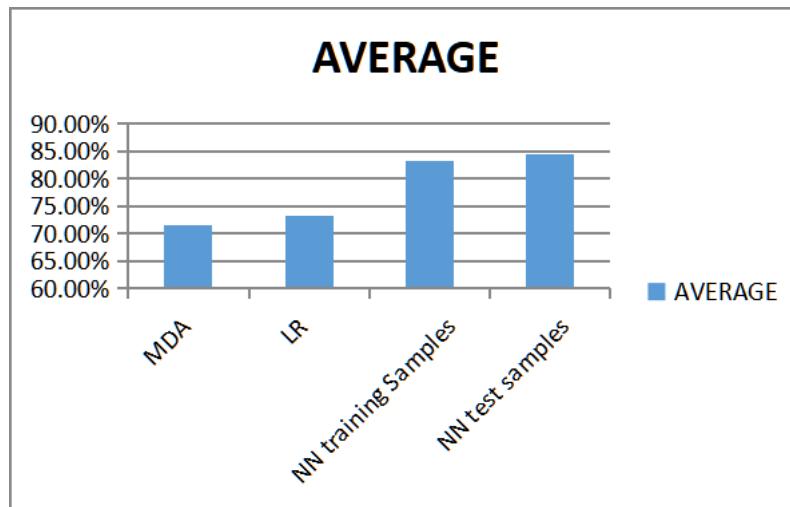

Figure 4. *Comparison of the Average Prediction Performance*

Conclusion

This study uses data taken from public companies financial statements, which shares listed on Indonesia stock exchange in the period of 2015 - 2012, taken from Indonesia Stock Exchange. MDA shows that the characteristics of companies that pay and not pay dividend are differentiated by price to book value, stock price, firm cycle, current ratio, ROA and Exchange rate. We used these variables as predictors both in LR and NN. By comparing the three metrics we conclude that neural network produce better model compare to MDA & LR. The average performance of Accuracy, Precision and Sensitivity in predicting wether the company pay or not to pay dividend is the best in Neural Network Model.

References

Abdou, H. A., Pointon, J., El-Masry, A., Olugbode, M., & Lister, R. J. (2012). A variable impact neural network analysis of dividend policies and share prices of transportation and related companies. *Journal of International Financial Markets, Institutions and Money*, *22*(4), 796–813. https://doi.org/10.1016/j.intfin.2012.04.008

Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance*, *18*(3), 505–529. https://doi.org/10.1016/0378-4266(94)90007-8

Bhattacharya, S. (1979). Imperfect Information, Dividend Policy, and "The Bird in the Hand" Fallacy. *The Bell Journal of Economics*, *10*(1), 259. https://doi.org/10.2307/3003330

Coulombe, D. (1985). Multiple discriminant analysis: A BASIC program for microcomputers. *Behavior Research Methods, Instruments, & Computers*, *17*(1), 135–136. https://doi.org/10.3758/bf03200914

Dorsey, R. E., Edminister, R. O., & Johnson, J. D. (1995). *Bankruptcy Prediction Using*

Artificial Neural Systems.

Franc-Dąbrowska, J., Mądra-Sawicka, M., & Ulrichs, M. (2020). Determinants of dividend payout decisions–the case of publicly quoted food industry enterprises operating in emerging markets. *Economic Research-Ekonomska Istrazivanja* , *33*(1), 1108–1129. https://doi.org/10.1080/1331677X.2019.1631201

Gordon, M. J. (1959). Dividends, Earnings, and Stock Prices. *The Review of Economics and Statistics*, *41*(2), 99–105.

Greene, W. H. (2000). *Econometric Analysis 4th Ed.*

Grullon, G., Michaely, R., & Swaminathan, B. (2002). Are Dividend Changes a Sign of Firm Maturity? *Journal of Business*, *75*(3), 387–424. https://doi.org/10.1086/339889

Ismiyanti, F. (2005). Uji Faktor-Faktor Pembeda Kebijakan Dividen : Analisis Diskriminan. *Manajemen & Bisnis*, *4*(1), 30–47.

Jensen, M. C., & Meckling, W. H. (1976). Theory of The Firm : Managerial Behavior, Agency Cost, and Ownership Structure. *Journal of Financial Economics*, *3*(July), 305-360. Retrieved from https://uclafinance.typepad.com/main/files/jensen_76.pdf

Laoh, L. C. (2019). Dividend Payout Forecast : Multiple Linear Regression vs Genetic Algorithm-Neural Network. *CogITo Smart Journal*, *5*(2), 252. https://doi.org/10.31154/cogito.v5i2.210.252-265

Lestari, H. S. (2018). Determinants of corporate dividend policy in Indonesia. *IOP Conference Series: Earth and Environmental Science*, *106*(1). https://doi.org/10.1088/1755-1315/106/1/012046

Lintner J. (1956). Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes. *The American Economic Review*, *46*(2), 97–113.

Miller, M., & Scholes, M. (1978). Dividends and Taxes. *Journal of Financial Economics*, *6*, 333–364.

Shefrin, H. M., & Thaler, R. H. (1988). the Behavioral Life-Cycle Hypothesis. *Economic Inquiry*, *26*(4), 609–643. https://doi.org/10.1111/j.1465-7295.1988.tb01520.x

Shefrin, H., & Statmant, M. (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence. *The Journal of Finance*, *40*(3), 777–790. https://doi.org/10.1111/j.1540-6261.1985.tb05002.x

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia Medica*, *24*(1), 12–18. https://doi.org/10.11613/BM.2014.003

Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica*, *26*(1), 24–36.