

Neural Network on Stock Prediction using the Stock Prices Feature and Indonesian Financial News Titles

Nur Ghaniaviyanto Ramadhan ¹, Imelda Atastina ^{1*}

¹*School of Computing, Telkom University
Bandung, Indonesia*

*imelda@telkomuniversity.ac.id

Abstract

Stocks are the most popular investments among entrepreneurs or other investors. When investing in stocks these investors tend to learn how to invest stocks correctly and the right conditions. For the problem of how to invest shares correctly can be used a variety of basic theories that already exist, but for the problem when the right time needs further learning. In this paper will purpose about stock price prediction using stock data indicators and financial titles data in Bahasa Indonesia. The machine learning model used is a multi-layer perceptron neural network (MLP-NN). MLP-NN is a feed-forward neural network consisting of a few neurons connected by connecting weights. The results obtained in this study by applying the MLP-NN model resulted in an accuracy of 80%.

Keywords: Stock Prices, Indonesian Financial News Title, Neural Network, Multi-Layer Perceptron

I. INTRODUCTION

Investment can be defined as the commitment of some money or other resources made today (present time) in the hope of obtaining benefits in the future [1]. Investments can be in the form of stocks, gold, deposits, mutual funds, or other securities. Stocks are the most popular investments among entrepreneurs or other investors. When investing in stocks these investors tend to learn how to invest stocks correctly and in the right conditions. The problem of how to invest shares correctly can be using a variety of basic theories that already exist, but for the problem when the right time needs further learning. Investors certainly do not want to suffer losses because they choose the wrong investment time, for example, which should today sell to make a profit but instead do buy which will add to the loss. Such a thing can certainly be predicted in advance with various machine learning models.

Some existing research discusses stock price predictions using machine learning models. Some call it the random walk theory [2] which mentions the stock price hypothesis is determined randomly. Surely if the share price is determined randomly will be very detrimental to investors. The current development of artificial intelligence implementation and the availability of stock price data can certainly make predictions much better than randomly [3]-[12]. In 1988, IBM predicted daily share price changes using NN [3]. Paper [4] tells about autoregressive models integrated moving average (ARIMA) with artificial NN (ANN) to predict time series and conduct comparative studies. The results [4] showed that JST is more beneficial in analyzing and processing nonlinen data. Paper [5] proposes an NN-based empirical methodology for designing automated trading systems

based on ANN. The paper [6] discusses stock price predictions using Twitter user mood data with the Self-Organizing Fuzzy Neural Network prediction method. Paper [7] discusses daily predictions of share prices using a combination of English news and stock indicators by applying the SVM model. Paper [8] tells the story of financial trading predictions using news stories and stock data, as well as implementing three different trading strategies by implementing the C.45 model. The paper [9] discusses the stock price based on hybrid models using the backpropagation neural network model. Paper [10] tells the story of predicting the amount of stock-based on financial news using the SVR model. Paper [11] discusses real dataset predictions of time series using ANN models. Paper [12] discusses market predictions using text mining techniques with headline data. Paper [13] discusses stock price predictions using technical indicators and headlines in English.

In addition to the machine learning models discussed in the paper [3]-[13] there are also deep learning models used for stock price predictions namely CNN [14]-[17], RNN [18],[19], and RCNN [20],[21],[22]. This study will predict the stock price by using a combination of stock history data and news headlines in Indonesian with the model used that is multi-layer perceptron neural network.

II. LITERATURE REVIEW

In paper [3]-[22] it certainly has shortcomings, such as in the word2vec process and stock data is done separately without merging the two datasets [7],[13]. In addition to paper [3]-[6],[8]-[12],[14]-[22] it uses only one type of data namely English-language news data or stock history data.

A. Previous Study

TABLE I
HIGHLIGHT CONTRIBUTIONS STUDY

Authors	Motivations	Merging Dataset	Method
Previous Study [3]-[6],[8]-[12],[14]-[22]	Stock History or English Language News	No	ANN, CNN, RNN, SVR, RCNN, F-NN
Previous Study [7],[13]	Stock History + English Language News	No	SVM, RNN, CNN
This Study	Stock History + Indonesian Language News	Yes	MLP-NN

Based on table I so, there is still the need for research that discusses two datasets, namely stock history data and Indonesian news titles, and the merger of the two datasets.

B. Neural network

Neural networks (NN) or rather artificial neural networks are branches of artificial intelligence. multi-layer perceptron form one type of neural network as depicted in taxonomy on Fig. 1 [23].

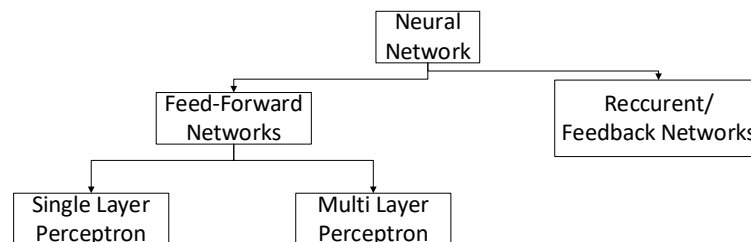


Fig. 1. Neural Networks Hierarchy

Multilayer Perceptron has been applied to a wide variety of tasks, all of which can be categorized as prediction, approximate function, or classification pattern. Predictions involve forecasting future trends in a time series of data given current and previous conditions. The function approach relates to modeling relationships between variables. Pattern classification involves classifying data into discrete classes. This study will use a multi-layer perceptron model using a forward-backward propagation algorithm. One of the models in NN is the multi-layer perceptron (MLP) which is used in this study because in this model the calculation process uses the weight value between networks and this model can solve very complex problems in the dataset.

III. RESEARCH METHOD

This paper will propose the process as in fig 2.

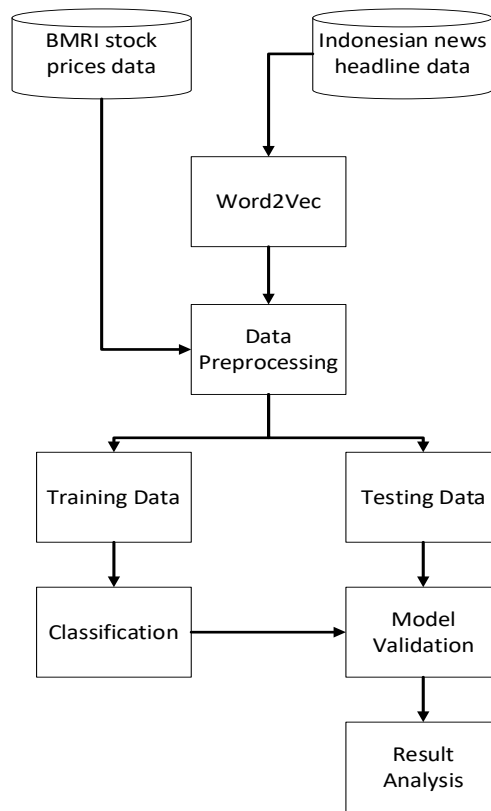


Fig. 2. Proposed Diagram

A. Dataset

In this process, collect datasets related to the share. This research, using a stock dataset from one of the banks in Indonesia, namely Bank Mandiri with the ticker code BMRI. The stock dataset is obtained through the yahoo finance website [24]. In addition to the dataset from the independent bank, this research also uses a dataset containing Indonesian-language news titles obtained through yahoo's website. The duration of the dataset BMRI and Indonesian language news titles used in the November 2020 period for each dataset. The number of datasets used in this research was considered representative for testing. For more detailed datasets can be seen in tables II and III.

TABLE II
DATASET BMRI

No	Attribute Name	Type Attributes
1	Date	Date
2	Open	Numeric
3	High	Numeric
4	Low	Numeric
5	Close	Numeric
6	Adj Close	Numeric
7	Volume	Numeric

For date information on the dataset, the date attribute depends on the stock exchange workday where if it is Saturday, Sunday, or a shared holiday then the exchange is off. The open attribute is the value that explains the stock price opened on the day. The high attribute is the value that explains the share price on the day reached its highest range on the day. The low attribute is the value that explains the share price on the day reached its lowest range on the day. Closed attributes explain the stock price closed on the day. The adj close attribute is an adjusted closing share price. The volume attribute is an indicator of crowded or quiet stock transactions on the day.

TABLE III
DATASET OF INDONESIAN NEWS TITLES

No	Attribute Name	Type Attributes
1	Date Indonesian news	Date
2	Title	Text

The dataset of Indonesian news titles used by taking five titles on the day of the active exchange to get a total of 105 rows of data. This dataset is based on text, in which the dataset for news titles will be converted to vector using word2vec before the classification process is carried out. The limitation of Indonesian news titles used in this study is the only news that belongs to the financial and economic category.

B. Word2Vec

At this stage will be done word2Vec process which will change the five lines of news titles in each day into vectors x and y . x and y values in the result word2vec interpret a vector 2D, so that if the value of x minus then he will be plotted to the left of the x -axis then on the contrary if the value plus he will be plotted to the right of the x -axis. The concept of word2vec is a cluster whereby measuring the cluster of word proximity to each other. So, it can be concluded that variable x in word2vec is an independent

variable, while variable y is a dependent variable. For examples of the top five rows results obtained in this process can be seen in Table IV.

TABLE IV
 RESULTS WORD2VEC

No	word	x	y
1	program	-83.11	74.57
2	restrukturisasi	31.53	-48.26
3	kredit	-41.92	-99.14
4	diperpanjang	-127.15	71.74
5	begini	149.41	28.22

In the word2vec process, the dataset of news titles in one day consists of five titles, which in the end will be calculated the average of the x and y variables for each of the five titles rows.

C. Data Preprocessing

In this process, there will be a merger of two datasets, namely the BMRI stock dataset and news titles result from a dataset that has been done word2vec process. The combined result dataset feature more clearly can be seen in Table V.

TABLE V
 COMBINED DATASET

No	Attribute Name	Type Attributes
1	Date	Date
2	Open	Numeric
3	High	Numeric
4	Low	Numeric
5	Close	Numeric
6	Adj Close	Numeric
7	Volume	Numeric
8	lenWr2Vec	Numeric
9	average X	Numeric
10	average Y	Numeric
11	Price	Nominal
12	Class	Nominal

The description of the price feature above is an interpretation of if the price on the day goes up then it is labeled 1 if the price goes down is labeled 0. Meanwhile, the description of the class feature above is an interpretation of if tomorrow's share price goes up (label 1) then now buy (label 0), on the contrary, if tomorrow's share price falls (label 0) then now sell (label 1). For lenwr2vec feature description is the vector length of the result of word2vec on each day, average X feature and average Y feature is the average value of vector x and y on each day.

Furthermore, feature selection is carried out which for the feature is not used, namely date, this is because the data used is clear in November 2020. The second feature that will not be used is Close, this is because there is already an Adj Close feature which is the result of closing shares that have been adjusted. BMRI stock price dataset and titles news before the merger have been balanced for data can be seen in Table VI.

TABLE VI
DATASET AFTER PREPROCESSING

Dataset	Number of Data
BMRI	21
Title News	21

D. Multi-Layer Perceptron Neural Network Classification

In this process, classification prediction will be done using a machine learning model namely Multi-Layer Perceptron Neural Network (MLP-NN). The experiment will be conducted using changes to the related parameters. After preprocessing, the features used in the prediction process amounted to 9 features. To split the percentage of the dataset using ratio training data 75% and testing data 25%.

TABLE VII
DATA SPLIT RATIO

Data Type	Number of Data	Percentage
Training Data	16	75%
Testing Data	5	25%
Total	21	100%

Split data conducted on this study randomly by the system and model validation using the built-in model from MLP-NN.

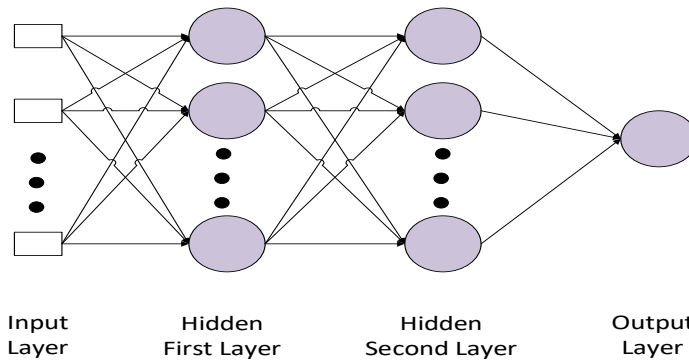


Fig. 3. Multi-Layer Perceptron Neural Network

Fig 3 is a neural network image used in this study. Generally, a neural network is divided into layers namely the input layer, the hidden layer, and the output layer. Each node in each layer has an error rate that will be used for the training process. A neural network with layers has the following concept.

1) *Input Layer*

This initial stage user input into each - each node, after each node in the input layer obtain the required data will be multiplied by the weight generates the amount or better known as an accumulator with the formula (1).

$$NET = O_1W_1 + O_2W_2 + \dots + O_n \quad (1)$$

2) *Hidden Layer*

At this stage discuss what happens in the inner layer. The relationship between nodes is associated with a value called weight. Each node must have its output, error, and weight. An error is a level of error contained in a node of the process performed. Weight is the weight of that node to another node on a different layer. Weight values range between -1 and 1. At this stage look for weight on each network that passes through the node, the formula used is (2).

$$W_n = \Sigma O_iW_i \quad (2)$$

To look for errors in the output layer can use a formula (3).

$$Error = Output \times ((1 - Output) \times (Target - Output)) \quad (3)$$

To look for errors in the next stage can use a formula (4).

$$Error = Output \times (1 - Output) \times W_n \quad (4)$$

As for looking for accumulators (5).

$$Acc = \Sigma W_n \times Error \quad (5)$$

Then look for a new weight (6).

$$W = W + (\alpha \times Output_{before} \times Error_{after}) \quad (6)$$

3) *Output Layer*

In the last stage, this is the output layer which is the output of a node. After the hidden layer process and then the formula (1) is inserted into the activation function used, the formula (7). The results obtained at this stage are binary guesses [0/1].

$$OUT = F(NET) \quad (7)$$

IV. RESULTS AND DISCUSSION

At this stage, experimentation is carried out by changing some parameters in the neural network. The scenario used is to increase or decrease the variables in NN such as hidden neurons, alpha, and epoch. The purpose of this first scenario is to find out which parameters on NN have the most influence. In addition to these scenarios, scenarios using word2vec and without word2vec were also carried out. The purpose of this second

scenario is to determine the effect of using a dataset of news titles. Table VIII can be seen for a more detailed scenario. To calculate the accuracy results obtained, formula (8) is used.

$$\text{Accuracy} = \frac{\text{True Value} - \text{Predictive Value}}{\text{True Value}} \tag{8}$$

Table VIII is an example of the form of predictive data results carried out.

TABLE VIII
EXAMPLES OF PREDICTED DATA RESULTS

Result of Testing	Result of Predict
Selling	Selling
Selling	Selling
Selling	Buying
Selling	Selling
Selling	Buying

TABLE IX
RESULT EXPERIMENTS WITH AND WITHOUT WORD2VEC

Hidden Neuron	Alpha	Max Epoch	Result of Testing with Word2Vec (%)	Result of Testing without Word2Vec (%)
100000	0.0001	100	60	40
10000	0.0001	100	80	20
10000	0.001	100	60	80
10000	0.001	10	40	20
100	0.1	100	40	60
1000	0.1	100	60	40
10000	0.1	100	80	0
100000	0.1	100	40	20
100000	0.01	100	80	40
10000	0.01	100	80	80
10000	0.2	100	60	20
10000	0.02	100	60	40
10000	0.5	100	60	40

Based on the experimental results in table IX the results obtained vary for each parameter. For parameters that cannot be changed i.e., neuron output, this is because the class used in the dataset is binary instead of multiclass. The epoch parameter is the rotation/iteration used, if the epoch gets bigger then the learning process

will be longer in terms of time if in terms of accuracy it is not much different from the epoch 100. The parameters that most affect the neural network algorithm are the number of hidden neurons and alpha. This can be seen in the results of experiments where if the number of hidden neurons more than 1000 has a testing accuracy of 80%. However, if the number of hidden layers is more than 1000 but the alpha value is greater than 0.1 the resulting accuracy value tends to decrease. If done experiments without using the results of word2vec then the results on the accuracy of testing obtained are much different even though the parameters are the same. Based on Table IX, the NN model can be adaptive to carry out the problem-solving process by considering a significant process, namely the hidden layer by calculating the value weights on each resulting network. In the experiment thirteen times, the results obtained with word2vec are more superior to without word2vec. This is because if only doing stock price predictions using stock history data is not enough, it is necessary to do a combination using a dataset of news titles. In this study, the titles influence the results of stock price predictions. So, the use of word2vec in Indonesian news titles can be said to influence stock prediction results.

V. CONCLUSION

Based on the proposed method that has been done, the conclusion that can be taken is a multi-layer perceptron neural network algorithm can predict the classification of stock data that has been combined with Indonesian news titles data. The results of word2vec influence stock price predictions. The most influential parameters on neural network algorithms are hidden neurons and alphas. The accuracy of testing produced in this experiment is the largest is 80%. For future work can be used on stock price prediction in other fields and use a more varied machine learning model.

REFERENCES

- [1] Tandelilin, Eduardus. "Dasar-dasar Manajemen Investasi." Taken from [http://repository. ut. ac. id/3823/1/EKMA5312-M1](http://repository.ut.ac.id/3823/1/EKMA5312-M1). Pdf, 2010. Accessed: Des, 10, 2020. (In Indonesian).
- [2] Malkiel, Burton G. "A random walk down Wall Street. W. W." (1973).
- [3] White, Halbert. "Economic prediction using neural networks: The case of IBM daily stock returns." *ICNN*. Vol. 2. 1988.
- [4] Zhang, G. Peter. "Time series forecasting using a hybrid ARIMA and neural network model." *Neurocomputing* 50 (2003): 159-175.
- [5] Vanstone, Bruce, and Gavin Finnie. "An empirical methodology for developing stockmarket trading systems using artificial neural networks." *Expert systems with Applications* 36.3 (2009): 6668-6680.
- [6] J. Bollen, H. Mao, and X. Zeng, Twitter mood predicts the stock market, *Journal of Computational Science*, 2(1), 1-8 (2011).
- [7] Y. Zhai, A. Hsu and S. Halgamuge, Combining news and technical indicators in Daily stock price trends prediction, *Advances in Neural Networks*, 1087-1096 (2007).
- [8] G. Rachlin, M. Last, D. Alberg and A. Kandel, ADMIRAL: A data mining based financial trading system, *IEEE Symposium on Computational Intelligence and Data Mining*, 720-725 (2007).
- [9] J. J. Wang, J. Z. Wang, Z. G. Zhang and S. P. Guo, Stock index forecasting based on a hybrid model, *Omega*, 40(6), 758-766 (2012).
- [10] R. P. Schumaker and H. Chen, A quantitative stock prediction system based on financial news, *Information Processing and Management*, 45(5), 571-583 (2009).
- [11] M. Khashei and M. Bijari, A novel hybridization of artificial neural networks and ARIMA models for time series forecasting, *Applied Soft Computing*, 11(2), 2664-2675 (2011).
- [12] A. K. Nassirtoussi, S. Aghabozorgi, T. Y Wah and D. C. L. Ngo, Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment, *Expert Systems with Applications*, 42(1), 306-324 (2015).
- [13] Vargas, Manuel R., et al. "Deep learning for stock market prediction using technical indicators and financial news articles." 2018 International Joint Conference on Neural Networks (IJCNN). IEEE, 2018.
- [14] Y. Kim, Convolutional neural networks for sentence classification, arXiv preprint arXiv:1408.5882, (2014).
- [15] X. Zhang and Y. LeCun, Text understanding from scratch, arXiv preprint arXiv:1502.01710, (2015).
- [16] P. Blunsom, E. Grefenstette and N. Kalchbrenner, A convolutional neural network for modelling sentences, *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, (2014).

- [17] R. Collobert and J. Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning, Proceedings of the 25th International Conference on Machine Learning, ICML, New York, NY, USA, 160–167 (2008).
- [18] T. Mikolov, M. Karafiát, L. Burget, J. Cernocký and S. Khudanpur, Recurrent neural network based language model, INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, September 26-30, 1045–1048 (2010).
- [19] O. Irsoy and C. Cardie, Opinion Mining with Deep Recurrent Neural Networks, EMNLP, 720-728 (2014).
- [20] C. Zhou, C. Sun, Z. Liu and F. Lau, A C-LSTM neural network for text classification, arXiv preprint arXiv:1511.08630, (2015).
- [21] S. Lai, L. Xu, K. Liu and J. Zhao, Recurrent Convolutional Neural Networks for Text Classification, AACL, 2267-2273 (2015).
- [22] Yu, Pengfei, and Xuesong Yan. "Stock price prediction based on deep neural networks." *Neural Computing and Applications* 32.6, 1609-1628, (2020).
- [23] Gardner, Matt W., and S. R. Dorling. "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences." *Atmospheric environment* 32, 14-15, 2627-2636, (1998).
- [24] Yahoo Finance, <https://finance.yahoo.com/quote/BMRI.JK/history?p=BMRI.JK>, Accessed December 10, 2020.
- [25] Yahoo, <https://id.search.yahoo.com/>, Accessed December 10, 2020.