Journal of Theoretical and Applied Information Technology
15th November 2022. Vol.100. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COMPARISON ARTIFICIAL NEURAL NETWORK METHODS OF BACKPROPAGATION AND LEARNING VECTOR QUANTIZATION FOR FORECASTING STOCK PRICES

SILVIA FIFI WARDHANI[1], DIYURMAN GEA[2]

[1,2]Information Systems Management Department, BINUS Graduate Program - Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia 11480

E-mail: [1] silvia.wardhani@binus.ac.id, [2] diyur@binus.edu

ABSTRACT

During the COVID-19 pandemic stock trading is a hot topic of discussion and encourages new investors which positive impact on the market modal. Shares of PT. XL Axiata, Tbk. (EXCL) was sluggish despite reporting a surge profit in 2021, this prompted research on how to predict the stock price of EXCL for attract investors and encourage company to be more active in carrying out business strategies. In recent years, Artificial Neural Networks (ANN) are quite used in macroeconomics forecasting, because of their ability to detect and relate linear and non-linear functions. In this study, two ANN methods were used to predict the stock price of EXCL with backpropagation (BP) and Learning Vector Quantization (LVQ). In the prediction results, model evaluation is needed to measure the forecasting model from both methods, resulting in the confusion matrix with the accuracy, sensitivity, and specificity are provided. This research is given some value to stock action suggestions at EXCL.

Keywords: *Stock Price EXCL, Artificial Neural Networks (ANN), Backpropagation (BP), Learning Vector Quantization (LVQ), Confusion Matrix*

1. INTRODUCTION

Many ways can be done by investors to invest, one of which is by investing in the capital market, the capital market is a place that trades various types of long-term financial instruments, one of the most preferred financial instruments by investors is stocks. Stocks are one of the most popular types of investment, it can be seen from the development of stocks from year to year which is increasing due to the high demand for stock investments. Shares are divided into several sectors based on the type of business, the shares that will be used in this analysis are shares in the telecommunications sector in Indonesia. Telecommunications stocks are one of the most sought-after stocks because these stocks can offer large returns for investors. This can happen because telecommunication service users in Indonesia continue to increase, it is known that the penetration of Indonesian internet users has increased from 64.8 percent in 2018 to 73.7 percent of the total population of Indonesia in mid-2020 [1]. The increase in users occurs because currently, almost everyone uses various telecommunications media, so the existence of telecommunications network provider services is very necessary. However, to become one of the leading telecommunications service providers is not an easy thing, because running this business requires a very large amount of capital to invest in infrastructure. To develop the telecommunications business, companies usually go public to get additional capital. A go public company or known as a public company is a company that has registered shares and is sold on the stock market. The sale of these shares is intended to obtain additional capital for the development of the company. Telecommunications stocks that have gone public are the three largest telecommunications companies in Indonesia, namely PT. Telekomunikasi Indonesia Tbk., PT. Indosat Tbk., and PT. XL Axiata Tbk. In this study, one of them was taken, namely PT. XL Axiata, Tbk (EXCL). EXCL recorded a good performance throughout 2021. The performance of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

EXCL in 2022 will still be supported by an increase in data demand.

Moreover, going forward, EXCL will develop 5G technology while continuing to strengthen its 4G network [2]. Shares of telecommunication issuer EXCL was sluggish despite reporting a surge in profits in 2021 [3] this prompted research on how to predict the stock price of EXCL in the next period to attract investors to continue to invest in the shares of this issuer and encourage companies to be more active in carrying out business plans or strategies that support market and technology developments in Indonesia. In addition, it is not only the company's internal factors that affect the demand for selling or buying shares. Therefore, this is something that must be considered by stakeholders to achieve future business targets to deal with fluctuating stock market conditions, which can go up or down as well as the prices of goods or commodities in the market [4]. In addition, investment activities do not only come from the finance department. Investments can also be made in the marketing division by opening new distribution networks, as well as in the production division by purchasing new equipment or machines (infrastructure) to increase production [5].

The stock market is often influenced by many fundamental factors and technical analysis. In this situation, data mining techniques can be one way to meet the demand for accurate predictions of financial variables. In the past, most predictive models were based on conventional statistical methods, such as time-series and multivariate analysis. The AI technique was developed to meet the increasing demand for tools and methods that can predict, detect, classify, summarize the structure of variables, and determine the relationships between them, without relying heavily on certain assumptions, such as linearity or error normal distribution. Many AI applications in the financial analysis have answered this demand by developing parametric non-linear models. In capital market research, stock prices or indices are notoriously difficult to predict using traditional forecasting methods, such as least squares regression due to the nature of the distribution of their data. As a result, several AI-based methodologies have been proposed which have better predictive ability than traditional models [6].

In this study, there are two methods to predict the stock price of EXCL in the market, namely Backpropagation (BP) and Learning Vector Quantization (LVQ). Based on the author's knowledge, no other studies are using these two methods in predicting the stock price of EXCL. In addition, this research was carried out during the COVID-19 pandemic, so the trend for stock trading is still a hot topic of discussion and encourages the emergence of new investors, especially retail investors, which can have a positive impact on the world capital market modal [7]. BP and LVQ methods are types of Neural networks used in data classification and carried out on time-series data. Time series data are the values of a variable that are sequential according to time, namely daily, monthly, or yearly. The data obtained will be used to create a time series model. The time series model is a forecast of future values based on past values that have been observed (and data collected) at regular time intervals [8]. Before forecasting, it is necessary to make a model that best fits the available parameters so that it has a value in the confusion matrix that is suitable for its classification. This model will be used to forecast the next period in the modelling of Artificial Neural Networks.

Neural Network shows an effective approach for general purposes to find out patterns, classification, grouping, and especially time series forecasting with a high degree of accuracy. So, this research help provide information to potential investors on which method is accurate in making decisions to buy or sell shares. And provide recommendations to investors in making choices when they want to buy or sell shares. The stock index price that is forecast for the suggested stock action is the closed price of the EXCL next period. The price that is usually predicted is the closing price of the next day. This is so that investors can find out the next price estimate as a consideration in deciding whether to keep it or sell it. The results of this study can be used as a contribution of thought for the company where this research can be used as input for company leaders as a basis for consideration for the future. And can improve decision-making which decision is taken from basic information and experience that is reviewed from various aspects.

In this study, questions that arise to identify the problems to be studied are as follows:

1. How to implement the method for forecasting stock action suggestions at EXCL with LVQ (Learning Vector Quantization) and Backpropagation.
2. From the two methods that produce the most accurate value in forecasting the EXCL stock action suggestions.

2. LITERATURE REVIEW

In this section, the author describes a literature review about neural network for forecasting stock price from related research

2.1 Previous Studies

In research conducted by [9], shares are one of the instruments traded in the capital market. Investing in stocks can also offer huge returns, but it is also risky. The analysis is needed in carrying out transactions on stocks, especially weekly stocks to reduce the risk of investing. From the experimental results of this system, it was found that the accuracy of the system for transaction actions was 72%. In addition, in research conducted by [10] artificial neural networks are one of the soft computing methods that are widely used and applied in various disciplines, including time-series data analysis. The main purpose of time series data analysis is to predict time-series data that can be widely used in various real time-series data, including stock price data. In this study, the author tries to apply the backpropagation neural network method to one of the indicators of stock price changes, namely the Composite Stock Price Index (JCI). The research was continued by calculating the level of accuracy and reliability of the methods that have been applied to the JCI data. This approach is expected to be an alternative way of forecasting JCI data as an indicator of changes in stock prices in Indonesia. In research [11] accurate automatic plant identification is needed to solve various problems in agriculture. In this study, a comparison of the performance of the Backpropagation and Learning Vector Quantization neural networks was carried out to classify plants based on leaf characteristics. The first step is to extract 31 features from leaves that represent shape, colour, and texture. Furthermore, Backpropagation and LVQ networks are used for leaf data classification. The evaluation process on Backpropagation produces a maximum accuracy of 0.952 while the maximum accuracy on the LVQ network is 0.420. LVQ's failure to get satisfactory results is caused by the high complexity of the data. Further research should be conducted to reduce the complexity of the data using techniques such as feature selection or feature transformation.

From previous research, there is a difference with this research, namely the forecasting method used, each using only one neural network method, either backpropagation or LVQ. Then the comparison data of the two methods used are EXCL daily stock price data which is fluctuating, not like leaf characteristics as previous studies which data is not trending.

2.2 Stock Analysis

Stock analysis conducted by investors generally uses two basic approaches, namely: technical analysis and fundamental analysis. Technical analysis attempts to estimate the price by observing changes in the stock price in the past. Technical analysis is an investment approach by studying data from stock prices and relating them to the trading volume that occurred and the economic conditions at that time. This analysis only considers price movements without paying attention to the performance of companies that issue shares. The price movement is related to events at that time such as economic influence, political influence, trade statement influence, psychological influence, and the influence of other issues [9]. Fundamental analysis is a factor that is closely related to the condition of the company, namely the condition of the organization's management of human resources and the company's financial condition which is related to the company's financial performance. Fundamental analysis tries to estimate stock prices in the future by estimating the value of fundamental factors that affect future prices and assigning these variables to obtain an estimated stock price [9].

2.3 Data Forecasting

Stock market predictions have always had a special fascination for researchers. While many scientific efforts have been made, no method has been found to accurately predict stock price movements. The difficulty of prediction lies in the complexity of modelling market dynamics. Even with the lack of consistent predictive methods, there have been some mild successes. Stock market research encapsulates two elemental trading philosophies: fundamentals and technical approaches. In fundamental analysis, stock market price movements are believed to come from relatively safe data [12]. Fundamentalists use numerical information such as earnings, ratios, and management effectiveness to determine forecasts for the future. In technical analysis, it is believed that market timing is the key. Technicians use charting and modelling techniques to identify price and volume trends. These people then rely on historical data to predict future outcomes.

2.4 Artificial Intelligence (Artificial Intelligence) and Artificial Neural Network (ANN)

Artificial intelligence (AI) may not have an agreed-upon definition, but someone who writes

about its history has definition in mind. According to [13], artificial intelligence is an activity aimed at making machines intelligent, and intelligence is a quality that enables an entity to function properly and with foresight in its environment. By that definition, many things — humans, animals, and some intelligent machines. AI is the ability of digital computers or computer-controlled robots to solve problems usually associated with higher human intellectual processing abilities. However, this definition also has weaknesses [14]. It will recognize for example that computers with a large memory that can store long texts and retrieve them on-demand display intelligent abilities, to memorize long texts certainly being considered a higher human intellectual processing capability, for example, the rapid multiplication of two 20-digit numbers.

According to [15], artificial neural networks are mathematical discoveries inspired by observations made in the study of biological systems, although they are loosely based on biology. An artificial neural network can be described as a mapping of input to output space. This concept is analogous to a mathematical function. The purpose of a neural network is to map inputs into desired outputs. Both biological systems, neurons, and feathers, serve a useful purpose, but the application of the principals involved has resulted in man-made discoveries that bear little resemblance to biological systems giving birth to creative processes. Artificial Neural Networks (ANN) is a biologically inspired form of distributed computing. It stimulates the function of a biological nervous system with a simple composition of interconnected elements (artificial neurons) operating in parallel.

2.5 LVQ (Learning Vector Quantization)

LVQ is a method of an artificial neural network created by Teuvo Kohonen. LVQ is an adaptive data classification method based on training data with the desired class information [16]. Although it is a supervised training method, LVQ uses unsupervised data clustering techniques to pre-process data sets and determine the cluster centre. The LVQ network architecture is almost like a competitive training network except that each output unit is associated with a particular class. A neural network (ANN) is one part of Artificial Intelligence (AI). It is an information processing paradigm inspired by the densely interconnected, parallel structure of mammalian information brain processes. The key element of the ANN paradigm to be distinguished from other methods of Artificial Intelligence is self-learning. Neural networks are also well known for their ability to process nonlinear problems. Ramazan Gençay in 1996 showed strong evidence of non-linear predictability found in stock market returns by using past buy and sell signals of the moving average rule. Therefore, ANN is applied in this study to increase the predictability of the stock market and to increase profitability. LVQ is a competition learning algorithm said to be a supervised version of the Self-Organizing Map Algorithm (SOM).

2.6 Backpropagation

The name backpropagation comes from a term used by Rosenblatt in 1962 for his attempt to generalize the perceptron learning algorithm to the multilayer case. There were many attempts to generalize perceptron learning procedures to several layers during the 1960s and 1970s, but none of them were very successful [17]. Backpropagation is a versatile tool for computer modelers, engineers, and cognitive scientists in general. The backpropagation network can be viewed as a non-parametric, nonlinear, structured statistical model. The architecture and its applications then demonstrate the successful implementation of algorithms for speech processing, fingerprint recognition, process control, etc. One of the most popular artificial neural network architectures is multilayer feedforward networks. In general, such a network consists of several neuron units as an input layer, one or more layers of hidden layer computational neuron nodes, and a layer of output computational neuron nodes. Error backpropagation is an MLPs algorithm that uses the principle of supervised learning.

3. RESEARCH METHODOLOGY

In this section, how to apply this methodology, how should understand the data itself, and how will do with it

3.1 Framework

By the formulation of the problem and research objectives, this study was structured with the focus of forecasting the stock value of EXCL in the future, using the Artificial Neural Network (ANN) technique. In recent years, ANN is quite widely used in forecasting in macroeconomics, because of its ability to detect and relate linear and non-linear functions through various variables. As with other forecasting techniques, to predict future stock values, the influencing variables must first be determined, based on previous studies and the availability of the data in Indonesia. It often happens in various previous studies, data that is only available in a certain period, or available in a different period,

will complicate the forecasting process and will affect the accuracy of the results. After getting the right variables, then these variables are used in forecasting techniques using the ANN method. So, in this study, two ANN methods were used to predict the stock price of EXCL with backpropagation (BP) and Learning Vector Quantization (LVQ) methods.

3.2 Data Analysis

The research data use stock data which are considered to have varied stock price movements so that they can represent the right action suggestions for investors. The author gets the stock price needed for research with a technical and fundamental approach from the historical price of one of the issuers on yahoo finance. And on the stock application made by the securities company PT Indo Premier Securities. The data used in this study is the stock price data of EXCL the last ten years from Yahoo! Finance with a daily frequency of early 2011 until the end of 2021 because during the last 10 years various events have occurred such as stocks have increased and slumped. Historical data from the share price of EXCL consists of six attributes or variables, namely date, opening price, highest price, lowest price, closing price, and transaction volume. The stock index price that is predicted for the suggested stock action is the closed price because this is intended so that investors can find out the next estimated closing price as a consideration in deciding whether to keep their shares or sell them. The total number of data obtained is 2734 rows.

3.3 Data Collection

This study uses time-series data, in which changes in current or future data are the results of price changes in the past. The research data use stock data which are considered to have varied stock price movements so that they can represent the right action suggestions for investors. The author gets the stock price needed for research with a technical and fundamental approach from the historical price of one of the issuers on yahoo finance. And on the stock application made by the securities company PT Indo Premier Securities.

3.4 Data Processing

The data that has been obtained will be subjected to a normalization process using the MIN-MAX Normalization method in equation (3.1), where z_i is the i-th variable of the z variable in the initial scale, z_{max} is the maximum value of the z variable, z_{min} is the smallest value of the z variable, and z_{inew} is the i-th value of the transformed value. The normalization process is carried out to simplify the computational process. The data will be divided into 2, namely training data and evaluation data. Training data is used for the learning process and test data is used for evaluation. The training data used is data from January 3, 2011, to April 23, 2020, or as much as 85%, while for the evaluation process using data from April 24, 2020, to December 30, 2021, or as much as 15% of the total data.

$$Z^i{}_{new} = \frac{Z_{max} - Z_i}{Z_{max} - Z_{min}} \quad (1)$$

3.5 How the System Works

In the training process, the input is entered first, this input is in the form of training data, then it will be normalized to reduce the value of the data so that the classification can be done better. This also applies to test data. In the normalization process, the data is converted or scaled to become data that has a range between 0 and 1. After the data is normalized, the data will become a feature that will be used in the classification process. Once the feature is obtained, then enter the input in the form of data labels. Once the required data (features and data labels) are complete, we can start the training process.

$$Value = \frac{(actual\ value - minimum\ value)}{(maximum\ value - minimum\ value)}\ or\ f(x) = \begin{cases} 0, & x < x\ average \\ 1, & x \geq x\ average \end{cases} \quad (2)$$

This normalization is done by combining the minimum and maximum values in the data.

In the testing process, first, enter the test data, then the data will be scaled or normalized to produce features. The process of scaling and normalization that occurs is the same as the normalization process during the training process. The feature will be matched into the model that has been obtained from the previous training results, the ANN method will process and determine which class is right for the data so that the feature will have a class in the result.

3.6 Forecasting Model Evaluation

A confusion matrix is a tool used for evaluating classification models to estimate the correct or incorrect object [18]. A matrix of predictions that will be compared with the original class of input or in other words contains information on the actual and predicted values of the classification. In the confusion matrix is a table that states the amount of test data that is correctly classified and the amount of test data that is incorrectly classified [19]. The confusion matrix is a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

method that is usually used to perform accurate calculations on the concept of data mining. This formula performs calculations with 4 outputs, namely: recall, precision, accuracy, and error rate [20].

Table 1: Confusion Matrix

		Prediction Class	
		True	False
Real Class	True	True Positive (TP)	False Positive (FP)
	False	False Negative (FN)	True Negative (TN)

Where:
1. TP is True Positive, namely the number of positive data classified correctly by the system.
2. TN is True Negative, that is, the number of negative data correctly classified by the system.
3. FN is False Negative, that is, the number of negative data classified incorrectly by the system.
4. FP is False Positive, namely the number of positive data but classified incorrectly by the system.

In other words, the accuracy value is a comparison between the data that is classified correctly and the whole data. The equation can obtain the accuracy value.

$$Accuracy = \frac{TP+TN}{(TP+TN+FN+FP)} \quad (3)$$

The precision value describes the amount of data in the positive category that is correctly classified divided by the total data that is classified as positive. The equation can obtain precision:

$$Precision = \frac{TP}{(TP+FP)} \quad (4)$$

Meanwhile, recall shows several comparisons of positive category data that are classified correctly by the system

$$Recall = \frac{TP}{(TP+FN)} \quad (5)$$

The error rate is a case that is identified incorrectly in some data, so it can be seen how big the error rate is in the system used. The error percentage can be calculated using the equation below:

$$Error\ rate = \frac{\text{FN+FP}}{(TP+TN+F\ \ +FP)} \quad (6)$$

4. RESULTS AND DISCUSSION

In this section, some of the results obtained in this study will be explained

4.1 Backpropagation Method

The following is the application of the Backpropagation algorithm Neural Network model to forecast the daily closing price of EXCL. Several stages of data preparation will be tested using the Backpropagation method, namely checking missing data, normalizing the data used, and dividing the data into two parts (data partitioning), namely data training and data testing.

Missing Data Check

From the results of the examination, it was found that there were no missing data or that the entire data was complete. The absence of missing data can be seen in each variable which shows 0 as the number of missing data.

Data Normalization

Data normalization is done to change the measurement scale of the original data into data form so that the data scale has the same value. In this study, Min-Max Normalization data using the formula Equation (1).

$$Z^{i}{}_{new} = \frac{Z_{max} - Z_i}{Z_{max} - Z_{min}}$$

Distribution Data Testing and Training

The division of data into two parts, namely where there is training data is called the training process and data testing is called the testing process. The results of the percentage distribution data used are as follows:

Table 2: Distribution Data Testing and Training

Distribution Data	percentage	Total Data
Train Dataset	85%	2,324
Test Dataset	15%	410
Total	100%	2,734

Data distribution training must have a greater proportion than data testing because for machine learning there are more events with training data patterns. This is useful when the engine generates a

Journal of Theoretical and Applied Information Technology
15th November 2022. Vol.100. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

model and feeds it to the test data to produce more accurate predictions of the testing data.

Application of Backpropagation Method to Predict Stock Action

Data categories are used for qualitative variables when the results of certain observations/measurements with one of the categories are known and can be calculated. Binary category example: Only two different results with two possible categories. More than 2 categories: When it has more than two possible outcomes. Data categories can be nominal (not sequential) or ordinal (sorted). Data closing price of shares EXCL after normalization needs to be given label action. In this research, it has 3 experiments have been doing as explained below

4.1.1 Experiment by assuming normal market conditions have no issues or trends are not fluctuating

```
f = ifelse(data[,'Close norm'] > 0.59,  "jual",
           ifelse(data[,'Close norm'] < 0.59, "beli", "tahan"))
f
View(f)
class(f)
```

This means that if the closing stock price is greater than 0.59, which is the average closing price data after normalization, it will be labelled “sell/jual”, whereas if the closing price is less than 0.59, it will be labelled “buy/beli” and then “hold/tahan” if both conditions are not met. So, in the distribution of data classes as a stock suggestion action that has been carried out, the data that was originally on a ratio scale became a nominal scale by dividing the data "sell", "buy", and "hold".

```
> table(test_label, predict(nn, test_, type = "class"))

test_label beli jual
      beli  161    0
      jual    0  249
```

Where on the results of stock action forecasting with the backpropagation method using data testing, it was found that if the test label "buy" was able to be predicted well with prediction results "buy" also as many as 161 data, while the test label "sell" could be predicted well with prediction results "sell" as many as 249 data.

4.1.2 Experiment by assuming the market is experiencing a pandemic storm and sluggish

```
#Flag action

f = ifelse(data[,'Close norm'] > 0.83,  "jual",
           ifelse(data[,'Close norm'] < 0.83, "beli", "tahan"))
f
View(f)
class(f)
```

This means that in the data if the closing stock price is greater than 0.83 which is the average closing price data for the pandemic period during 2020 after normalization, it will be labelled “sell”, whereas if the closing price is less than 0.83 it will be labelled “buy”. then do "hold" if it does not meet the two conditions. During the 2020 coronavirus outbreak, the most visible impact of the COVID-19 crisis on financial markets was its effect on global stock markets [21].

```
> table(test_label, predict)
          predict
test_label beli jual
      beli  359    0
      jual    0   51
```

Where on the results of stock action forecasting with the backpropagation method using data testing, it is found that if the test label "buy" can be predicted well with prediction results "buy" also as much as 359 data, while the test label "sell" can be predicted well with prediction results "sell". "a total of 51 data.

4.1.3 Experiment by assuming the market condition is very good and the trend continues to increase

```
#Flag action

f = ifelse(data[,'Close norm'] > 0.35,  "jual",
           ifelse(data[,'Close norm'] < 0.35, "beli", "tahan"))
f
View(f)
class(f)
```

This means that if the closing stock price is greater than 0.35, which is the average closing price data for 2014 (it is assumed that EXCL shares are positive in this period) after normalization, it will be labelled "sell", while on the contrary if the closing price is lower of 0.35, it is labelled "buy" then "hold" is carried out if it does not meet the two conditions.

```
> table(test_label, predict)
          predict
test_label beli jual
      beli   63    0
      jual    0  347
```

Journal of Theoretical and Applied Information Technology
15th November 2022. Vol.100. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Where in the results of stock action forecasting with the backpropagation method using data testing, it is found that if the test label "buy" can be predicted well with predictive results "buy" also as much as 63 data, while the test label "sell" can be predicted well with predictive results "sell". as many as 347 data.

4.2 Learning Vector Quantization (LVQ) Method

The following is the application of the Neural Network model with the Learning Vector Quantization (LVQ) algorithm to forecast the daily closing price of EXCL. Several stages of data preparation will be tested using the Learning Vector Quantization (LVQ) method, namely checking missing data, normalizing the data used, and dividing the data into two parts (data partition), namely training data, and testing data.

Missing Data Check

It is necessary to check for missing data before the research continues with other stages such as extracting features and building models, the Learning Vector Quantization (LVQ) method itself also checks for missing data. From the results of the examination in the study, it is known that there is no missing data or that all the data is complete. The absence of missing data can be seen in each variable which shows 0 as the number of missing data.

Data Normalization

Data normalization is done to change the measurement scale of the original data into data form so that the data scale has the same value. In this study, Min-Max Normalization data using the formula Equation (1).

$$Z^{i}{}_{new} = \frac{Z_{max} - Z_i}{Z_{max} - Z_{min}}$$

Distribution Data Testing and Training

Categorical data is used for qualitative variables when the results of certain observations/measurements with one of several categories are known and can be calculated. The division of the data into two parts, namely where there is training data, which is called the training process and the test data is called the testing process. While the process of testing is the accuracy of the model obtained from the training process. The results of the distribution of the percentage distribution of the data used are as follows:

Table 3: Distribution Data Testing and Training

Distribution Data	Percentage	Total Data
Train Dataset	85%	2,324
Test Dataset	15%	410
Total	100%	2,734

The percentage of training data distribution must be greater than the testing data because machine learning is more trained with training data patterns. This is useful when the engine generates a model and feeds it to the test data to produce more accurate predictions of the testing data.

Application of LVQ Method to Predict Stock Action

Data closing price of shares EXCL after normalization needs to be given label action. In this research, it has 3 experiments have been doing as explained below

4.2.1 Experiment by assuming that the normal market conditions have no issues, or the trend is not fluctuating

```
#Pembagian kelas
f = ifelse(data[,'Close norm'] > 0.59,  "jual",
           ifelse(data[,'Close norm'] < 0.59, "beli", "tahan"))
f
View(f)
class(f)
```

This means that if the closing stock price is greater than 0.59, which is the average closing price data after normalization, it will be labelled “sell”, whereas if the closing price is less than 0.59, it will be labelled “buy” and then “hold”. if both conditions are not met.

```
> table(test_label, predict)
          predict
test_label beli jual
      beli  180    0
      jual    0  230
```

Where on the results of stock action forecasting with the Learning Vector Quantization (LVQ) method using testing data, it is found that if the test label "buy" can be predicted well with prediction results "buy" also as much as 180 data, while the test label

"sell" can be predicted well with the prediction results of "selling" as much as 230 data.

4.2.2 Experiment by assuming the market is experiencing a pandemic storm and sluggish

```
f = ifelse(data[,'Close norm'] > 0.83,  "jual",
           ifelse(data[,'Close norm'] < 0.83, "beli", "tahan"))
f
View(f)
class(f)
```

This means that in the data if the closing stock price is greater than 0.83 which is the average closing price data during the pandemic period which causes the stock to sluggish and the market condition is also uncertain after normalization, it is labelled "sell", while on the contrary if the closing price is smaller of 0.83 then it is given the label "buy" then "hold" is carried out if it does not meet the two conditions.

```
> table(test_label, predict)
          predict
test_label beli jual
      beli  187  174
      jual    0   49
```

Where on the results of stock action forecasting with the Learning Vector Quantization (LVQ) method using testing data, it is found that if the test label "buy" can be predicted well with prediction results "buy" also as many as 187 data, while the test label "sell" can be predicted well with the prediction results "sell" as much as 49 data. However, there is a prediction error of 174 data, where the test label is "sell" but is predicted to be "buy".

4.2.3 Experiment by assuming the market condition is very good and the trend continues to increase

```
f = ifelse(data[,'Close norm'] > 0.35,  "jual",
           ifelse(data[,'Close norm'] < 0.35, "beli", "tahan"))
f
View(f)
class(f)
```

This means that in the data if the closing stock price is greater than 0.35 which is the average closing price data when the stock condition improves in the market after normalization, it is labelled "sell", while on the contrary if the closing price is less than 0.35 then it is labelled "buy". then "hold" if it does not meet both conditions.

```
> table(test_label, predict)
          predict
test_label beli jual
      beli   84    0
      jual  100  226
```

Where on the results of stock action forecasting with the Learning Vector Quantization (LVQ) method using testing data, it is found that if the "buy" label test can be predicted well with the "buy" prediction results as well as 84 data, while the "sell" label test can be predicted well with the prediction results of "selling" as much as 226 data. However, there is a prediction error of 100 data, where the test label is "buy" but is predicted to be "sell".

4.3 Comparison of Forecasting Results

In this section will be explained about comparison of result experiment

4.3.1 Experiment by assuming that the normal market conditions have no issues, or the trend is not fluctuating

Table 4: Comparison of Forecasting Results

Confusion Matrix Measurement	Backpropagation	LVQ
Accuracy	1	1
No Information Rate	0.6073	0.561
Sensitivity	1	1
Specificity	1	1
Pos Pred Value	1	1
Neg Pred Value	1	1
Prevalence	0.3927	0.439
Detection Rate	0.3927	0.439
Detection Prevalence	0.3927	0.439
Balanced Accuracy	1	1

When market conditions are normal or the trend is not fluctuating, it produces Accuracy, Sensitivity, Specificity, Pos Pred Value, Neg Pred Value, and Balanced Accuracy with the same value. While the No Information Rate produces a value in the backpropagation method that is greater than the LVQ method. Prevalence, Detection Rate, and Detection Prevalence produce a value in the LVQ method that is greater than the backpropagation method.

4.3.2 Experiment by assuming the market is experiencing a pandemic storm and sluggish

Table 5: Comparison of Forecasting Results

Confusion Matrix Measurement	Backpropagation	LVQ
Accuracy	1	0.5756
No Information Rate	0.8756	0.5439
Sensitivity	1	1
Specificity	1	0.2197
Pos Pred Value	1	0.5180
Neg Pred Value	1	1
Prevalence	0.8756	0.4561
Detection Rate	0.8756	0.4561
Detection Prevalence	0.8756	0.8805
Balanced Accuracy	1	0.6099

When the situation is affected by the 2020 pandemic storm and assumes that the market is in a sluggish state, it produces Sensitivity and Neg Pred Value with the same value. Meanwhile, Accuracy, Specificity, No Information Rate, Pos Pred Value, Prevalence, Detection Rate, and Balanced Accuracy produce values in the backpropagation method that are greater than the LVQ method. However, for Detection Prevalence, the value in the LVQ method is greater than the backpropagation method.

4.3.3 Experiment by assuming the market condition is very good and the trend continues to increase

Table 6: Comparison of Forecasting Results

Confusion Matrix Measurement	Backpropagation	LVQ
Accuracy	1	0.7561
No Information Rate	0.8463	0.5512
Sensitivity	1	0.4565
Specificity	1	1
Pos Pred Value	1	1
Neg Pred Value	1	0.6933
Prevalence	0.1537	0.4488
Detection Rate	0.1537	0.2049
Detection Prevalence	0.1537	0.2049
Balanced Accuracy	1	0.7283

When market conditions are improving and the trend continues to increase, it produces Specificity and Post Pred Value with the same value. Meanwhile, Accuracy, Sensitivity, No Information Rate, Neg Pred Value, and Balanced Accuracy produce a higher value for the backpropagation method than the LVQ method. However, Prevalence, Detection Rate, and Detection Prevalence produce a value in the LVQ method that is greater than the backpropagation method.

From these three tests, it is found that the confusion matrix values such as Accuracy and balance accuracy of the backpropagation method are better than the LVQ method when market conditions are considered not in normal conditions such as a pandemic in 2020 and business conditions are growing in 2014. Assuming market conditions are moderate normal then the value of accuracy, sensitivity, and specificity of the two methods is the same. So, it can be concluded for stock forecasting EXCL by using the neural network method is influenced by the external state of the data itself.

5. CONCLUSIONS AND RECOMMENDATIONS

5. 1 Conclusion

1. The value of the confusion matrix such as Accuracy and balance accuracy of the backpropagation method is better than the LVQ method when market conditions are considered not in normal conditions such as a pandemic in 2020 and business conditions are growing in 2014. Assuming market conditions are normal, the value of accuracy, sensitivity, and specificity of the two methods is the same. So, it can be concluded for stock forecasting EXCL by using the neural network method is influenced by the external state of the data itself.

2. By using the neural network method, especially the backpropagation and LVQ methods, can provide more accurate stock action suggestions for investors and help companies make decisions to continue growing their business in market uncertainty.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5.2 Suggestions

Suggestions put forward after this research takes place include:

1. It is necessary to have other variables that are assumed to affect stock prices in the market as a consideration for making decisions and predicting stock prices in the future.

2. Can use neural network methods or other artificial intelligence to predict a company's stock.

3. Further research can use statistical software or other more accurate and up-to-date statistical or business intelligence tools by market and business needs.

REFERENCES:

[1] Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), "Sudah Selektifkah Anda dalam Memilih Penyedia Jasa Internet?", Security Advisory. Apr 14, 2021. [Online], Available: "https://bssn.go.id/sudah-selektifkah-anda-dalam-memilih-penyedia-jasa-internet/". [Accessed July. 22, 2022]

[2] N. Qolbi & A. P. Perwitasari, "Ini Rekomendasi Saham dan Prospek XL Axiata (EXCL) Usai Cetak Kinerja Ciamik di 2021". Feb 22, 2022. [Online], Available: "https://investasi.kontan.co.id/news/ini-rekomendasi-saham-dan-prospek-xl-axiata-excl-usai-cetak-kinerja-ciamik-di-2021". [Accessed July. 22, 2022]

[3] A. K. Saumi & Hafiyyan, "Laba Xl Axiata Melesat pada 2021, Saham EXCL Malah Lesu". Feb 21, 2022. [Online], Available: "https://market.bisnis.com/read/20220221/192/1502856/laba-xl-axiata-melesat-pada-2021-saham-excl-malah-lesu". [Accessed July. 22, 2022]

[4] Otoritas Jasa Keuangan (OJK), "Penyebab Naik Turun Harga Saham Suatu Perusahaan". [Online], Available: "https://sikapiuangmu.ojk.go.id/FrontEnd/CMS/Article/10507". [Accessed July. 22, 2022]

[5] Prasetyo, Guno. *Prediksi Harga Saham Sektor Keuangan dan Sektor Infrastruktur di Indonesia dengan Model ARIMA*. Diss. Universitas Brawijaya, 2018.

[6] Chen, Wun-Hua, Jen-Ying Shih, and Soushan Wu. "Comparison of support-vector machines and back propagation neural networks in forecasting the six major Asian stock markets." *International Journal of Electronic Finance* 1.1, 2006, 49-67.

[7] S. R. Fadly, "Aktivitas Pasar Modal Indonesia Di Era Pandemi". Mar, 31 2021 . [Online], Available: "https://www.djkn.kemenkeu.go.id/kpknl-kupang/baca-artikel/13817/Aktivitas-Pasar-Modal-Indonesia-Di-Era-Pandemi.html". [Accessed July. 22, 2022].

[8] Pham, Hoang, ed. *Springer handbook of engineering statistics*. Vol. 49. London: Springer, 2006.

[9] I. M. G. Inggrayana, W. Widodo, and L. A. Hermanto. "Saran Aksi Saham Dengan Pendekatan Fundamental Dan Teknikal Menggunakan Metode Learning Vector Quantization Neural Network." *INTEGER: Journal of Information Technology* 1.2, 2017.

[10] S. Hansun. "Peramalan Data IHSG Menggunakan Metode Backpropagation." *Ultimatics: Jurnal Teknik Informatika* 5.1, 2013, 26-30.

[11] S. Adinugroho and Y. A. Sari. "Perbandingan Jaringan Learning Vector Quantization dan Backpropagation pada Klasifikasi Daun Berbasiskan Fitur Gabungan." *Jurnal Informatika dan Multimedia* 9.2, 2018, 58-64.

[12] R. Vivek and S. Bobde. "Stock market forecasting techniques: literature survey." *International Journal of Computer Science and Mobile Computing* 5.6, 2016, 500-506.

[13] N. J. Nilsson. *The quest for artificial intelligence*. Cambridge University Press, 2009.

[14] E. Wolfgang. *Introduction to artificial intelligence*. Springer, 2018.

[15] Priddy KL, Keller PE. *Artificial neural networks: an introduction*. SPIE press, 2005.

[16] J. R. Jang, C.T. Sun, and E. Mizutani. "Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]." *IEEE Transactions on automatic control* 42.10, 1997, 1482-1484.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[17] Y. Chauvin, and D. E. Rumelhart. *Backpropagation: theory, architectures, and applications*. Psychology press, 2013.

[18] I. Nawangsih and S. Fauziah. "Prediksi Pengangkatan Karyawan dengan Metode Algoritma C5. 0 (Studi Kasus PT. Mataram Cakra Buana Agung)." *Pelita Teknologi* 16.2, 2021, 24-33.

[19] A. Indriani. "Klasifikasi Data Forum dengan menggunakan Metode NaÃ¯ ve Bayes Classifier." *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*. Vol. 1. No. 1, 2014.

[20] S. H. Waluyo and P. Prihandoko. "Klasifikasi Pemanfaat Program Beras Sejahtera (RASTRA) Berdasarkan Tingkat Kemiskinan Dengan Menggunakan Algoritma Decision Tree C4. 5 Berbasis Particle Swarm Optimization." *Energy-Jurnal Ilmiah Ilmu-Ilmu Teknik* 7.2, 2017, 19-24.

[21] Ozili PK, Arun T. Spillover of COVID-19: impact on the Global Economy. Available at SSRN 3562570, 2020 Mar 27.