

Intervention Analysis and Machine Learning to Evaluate the Impact of COVID-19 on Stock Prices

Hendri Prabowo^{1*}, Iman Rais Afandy¹

¹Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

*Corresponding author: hendriprabowo88@gmail.com

Received: 4 February 2021

Accepted: 27 March 2021

Published: 30 March 2021

ABSTRACT – The purpose of this study is to evaluate the impact of the COVID-19 outbreak on composite and individual stock prices in China, the USA, South Korea, and Indonesia by using an intervention model and comparing the results of its predictions with a machine learning model, i.e. neural network (NN) and deep learning neural network (DLNN). This intervention model can be used not only to find out the magnitude of the effect of COVID-19 on the stock price, but also the period of the effect. The composite stock price data used are KS11, 000001.SS, DJI, and JKSE, while the individual stock price data used are TLKM and EXCL. The data used is the daily stock data. The analysis shows that COVID-19 hurts stock prices both in countries that have passed the peak period and are still in the peak period of COVID-19. The impact is not directly after the first case of COVID-19 in each country. The lowest stock price occurred at the end of March 2020 in each country. Different conditions were shown by individual stock prices in the telecommunications sector that showed a positive trend after the end of April 2020. Generally, for all stock prices, intervention models are better for forecasting in-sample data and explanation impact COVID-19 on stock price, whereas machine learning models are better for forecasting out-of-sample data.

Keywords– COVID-19, Intervention, Machine Learning, Stock Price.

I. INTRODUCTION

Stock prices in the world are greatly influenced by important events [1]. Many important events in the world that affect stock prices, such as the occurrence of disasters [2], political events [3], and environmental changes [4]. Besides, stock prices in the world are also greatly affected by the presence of virus outbreaks that spread in several countries (epidemic diseases). Some epidemic diseases that affect stock prices are severe acute respiratory syndrome (SARS) in 2003 [5] [6], bird flu (H5N1) in 2005 [7], and Ebola in 2014 [8]. In general, the existence of an epidemic is very influential on the shake of the economic sector including stock prices.

At the end of 2019, a new pandemic called COVID-19 appeared in China. If the epidemic only spreads in a few countries, COVID-19 has spread to almost all countries. COVID-19 is a virus that causes diseases ranging from flu to various diseases [9]. The existence of COVID-19 can greatly affect various sectors, including stock prices. So, it is very necessary to analyze the impact of COVID-19 on stock prices.

Previous research on analyzing the impact of COVID-19 on stock prices has been carried out. [10] analyzed the impact of COVID-19 on the composite stock prices of China (000001.SS) and the United States (DJI) using a simple regression method. The results of his study showed that the addition of a positive case of COVID-19 reduced the composite stock prices in both countries. [11] analyzed the impact of COVID-19 on composite stock prices in China using panel data regression. The analysis showed that the number of positive cases and the number of daily deaths hurt the daily stock price.

One method that can be used to analyze the impact of an event such as COVID-19 is a time series approach with intervention analysis. Intervention analysis is a special type of time series model used to evaluate the impact of external factors such as disasters and internal factors such as political policy. Using intervention analysis, the magnitude and duration effect of the intervention will be known [12]. There have been many studies in various fields using intervention analysis. Box and Tiao used an intervention model for economic and environmental problems [13]. Vujic et al. used an intervention model to analyze the impact of changing political conditions in Virginia on the number of criminal cases [14].

This study uses intervention analysis to evaluate the impact of COVID-19 on composite stock prices in various countries (KS11, 000001.SS, DJI, and JKSE) and individual stock prices in Indonesia (TLKM and EXCL). The forecast results of the intervention analysis will be compared with machine learning methods, i.e. neural networks (NN) and deep learning neural networks (DLNN). In general, the contents in this paper are: Section 2 discusses the methods and dataset used in the analysis; Section 3 is the result of the analysis; Section 5 is the discussion; and Section 6 is the conclusion.

II. MATERIAL AND METHODS

A. Autoregressive Integrated Moving Average

The autoregressive integrated moving average (ARIMA) is one of the popular methods in time series analysis. The ARIMA model (p, d, q) is a combined model of AR and MA after differencing order d. In determining the AR and MA order, the data must require a stationary condition. ARIMA modeling can be done with the Box-Jenkins procedure, i.e. identification, parameter estimation, diagnostic testing, and forecasting [15]. In general, the ARIMA model (p, d, q) can be written as follows [12]:

$$\phi_p(B)(1-B)^d Y_t = \theta_0 + \theta_q(B)\alpha_t \tag{1}$$

where $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ and $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$.

B. Intervention Analysis

Intervention analysis is a special type of time series model used to evaluate the influence of external factors such as disasters and internal factors such as political policy, especially the magnitude and duration of the effects. There are two types of intervention models, i.e. the pulse function and the step function [12]. In general, the intervention model can be written as follows [12] [16]:

$$Y_t = \frac{\omega_s(B)B^b}{\delta_r(B)} X_t + N_t \tag{2}$$

where,

$$N_t = \frac{\theta_q(B)}{\phi_p(B)(1-B)^d} \alpha_t \tag{3}$$

where Y_t is the response variable at time t and X_t is a binary variable that indicates the presence or absence of intervention at time t. X_t can be a step function or pulse function. Then, known $\omega_s(B) = (\omega_0 - \omega_1 B - \dots - \omega_s B^s)$ and $\delta_r(B) = (1 - \delta_1 B - \dots - \delta_r B^r)$.

Based on equation (2) shows the value and period of the intervention effect based on the order b, s, and r. Order b shows the start of the intervention effect, order s shows the time needed for the intervention effect to be stable, and order r shows the pattern of intervention. Equation (4) shows the impact of the intervention model on time series data.

$$Y_t^* = \frac{\omega_s(B)B^b}{\delta_r(B)} X_t \tag{4}$$

The step function is a type of intervention that lasts long term and can be written in equation (5). The pulse function is a type of intervention that occurs at time T and can be written in equation (6).

$$S_t = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases} \tag{5}$$

$$P_t = \begin{cases} 0, & t = T \\ 1, & t \neq T \end{cases} \tag{6}$$

where T is the time the intervention starts. If the step function has the order b = 1, s = 2 and r = 0, we get the equation of the impact of the intervention model as follows:

$$Y_t^* = \omega_0 S_t - \omega_1 S_{t-1} - \omega_2 S_{t-2} \tag{7}$$

the form of the impact of the intervention on equation (7) can also be written as follows:

$$S_t = \begin{cases} 0, & t < T \\ \omega_0, & t = T \\ \omega_0 - \omega_1, & t = T + 1 \\ \omega_0 - \omega_1 - \omega_1, & t \geq T + 2 \end{cases} \tag{8}$$

In modeling with a single input intervention model, there are generally 2 steps. Previously the data was divided into data before the intervention (t = 1, 2, ... T-1) and data after the intervention (t = T, T + 1, ..., n). The following are steps in modeling the intervention:

1st Step

1. Apply the ARIMA model on data before the intervention (N_t).
2. Forecasting all data using the ARIMA model on the data before and after the intervention.
3. Calculating the value of the response function (standardized residual)

$$\lambda_t = \frac{(Y_t - \hat{Y}_t)}{\sigma_a} \tag{9}$$

where Y_t is actual data, \hat{Y}_t is forecast data from ARIMA model and σ_a is root mean square error (RMSE) value from ARIMA model.

4. Make a plot / control chart response function with a limit of ± 3 .

2nd Step

1. Identifying the order b, r, and s based on the response function plot.
2. Estimating the parameters of the intervention model, checking the significance of the parameters, diagnostic tests (white noise and normal distribution) residuals and forecasting based on the intervention model.

C. Neural Network

A neural network (NN) is a machine learning method that can be used to predict time series data [17] [18] [19]. In general, the most commonly used NN architecture is the feed-forward neural network (FFNN). The structure of FFNN consists of several layers, i.e. one input layer, one or more hidden layers, and one output layer. Each layer consists of one or more neurons. Where each neuron will receive information from the previous neuron [20]. Here is a general equation from NN:

$$\hat{Y}_{(t)} = f^o[\sum_{j=1}^q [w_j^o f_j^h (\sum_{i=1}^p w_{ji}^h X_{i(t)} + b_j^h) + b_j^o]], \tag{10}$$

where, b^o, b_j^h ($j=1, \dots, q$), w_{ji}^h ($j=1, \dots, q; i=1, \dots, p$), w_j^o ($j=1, \dots, q$) is a parameter of the NN model, f^o is an activation function at the output layer, f_j^h ($j=1, \dots, q$) is an activation function at the hidden layer, p is number of input variables and q is number of neurons in hidden layer. In the NN model, if there is one hidden layer, it is called the NN or FFNN model, and if there are two or more hidden layers, it is called a deep learning neural network (DLNN). The architecture of NN and DLNN is shown in Figure 1.

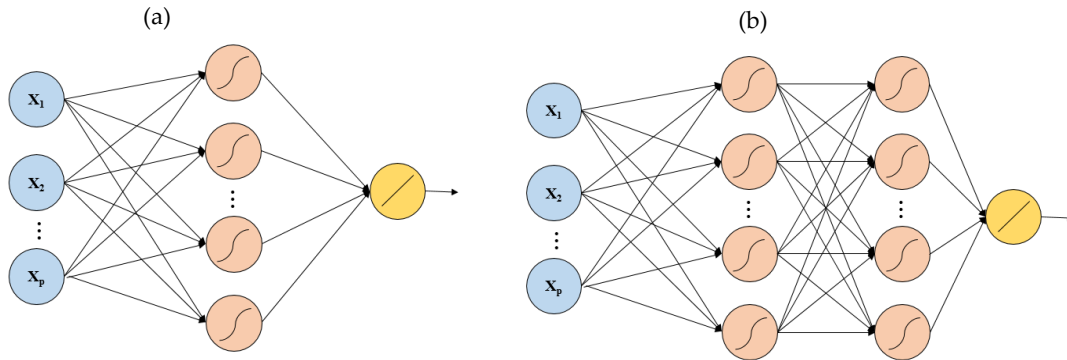


Figure 1. NN architecture with p neurons in the input layer (a) and DLNN architecture with p neurons in the input layer and 2 hidden layers (b)

D. Evaluation of Performance

In forecasting, it is very important to evaluate the results of the forecast to determine the method that produces the smallest error. One measure used to evaluate the forecast results from the time series method is the root mean square error (RMSE). The formula of RMSE is as follows [12]:

$$RMSE = \sqrt{\sum_{t=1}^n \frac{(Y_t - \hat{Y}_t)^2}{n}} \tag{11}$$

where Y_t is actual data, \hat{Y}_t is forecast data and n is the number of data.

E. Dataset

The data used in this study is data about composite stock prices and individual stock prices. The data used is daily stock data from 9 January 2019 to 20 May 2020. Daily stock data from 9 January 2019 to 8 May 2020 is used as in-sample data, and daily stock data from 11 May 2020 to 20 May 2020 is used as out-of-sample data. In-sample data is used to form a model, while out-of-sample data is used to evaluate forecast results. The data used is sourced from <https://finance.yahoo.com/>. Data on composite stock prices are used in four countries, namely Indonesia, South Korea, China, and the United States. Two individual stock price data in Indonesia for the infrastructure, utilities, and transportation sectors are also used. Can be seen in Table 1 for the names of stock prices used in this study.

Table 1. Stock Prices

Variable	Stock Prices Name
Composite	
Y_1	KOSPI Composite Index (KS11) – South Korea
Y_2	SSE Composite Index (000001.SS) – China
Y_3	Dow Jones Industrial Average (DJI) – USA
Y_4	IDX Composite Index (JKSE) – Indonesia
Individual	
Y_5	PT. Telekomunikasi Indonesia (TLKM)
Y_6	PT. XL Axiata (EXCL)

III. RESULTS

A. Stock Market Conditions

The analysis conducted begins with data visualization. Based on data from <https://www.worldometers.info/>, almost all countries have been affected by the COVID-19 virus. Some countries have passed the peak period, but also some countries are still in the peak period. Figure 2 shows the time series of cumulative case plots in Indonesia, South Korea, China, and the United States from January 22, 2020, to May 8, 2020. It can be seen that until May 8, 2020, the COVID-19 cumulative number curve in Indonesia and the United States is still showing an upward trend. It means that the increase in the number of daily cases is still very high, so it can be seen that Indonesia and the United States are still in the peak period of the spread of COVID-19. As for South Korea and China, it can be seen that the cumulative case curve has leveled off. It means that the addition of positive cases daily tends to be slightly even 0. So, it can be seen that South Korea and China have passed the peak period of the spread of the COVID-19 virus.

From Figure 2, it can be seen that COVID-19 is still spread in these four countries. The impact of COVID-19 is very broad in the health sector, economic, social, and cultural aspects. COVID-19 has been designated as a pandemic, meaning that it has spread throughout the world. The spread of disease in the category of endemic and pandemic greatly affects the condition of stock prices in the world. Endemic diseases such as SARS in 2003 greatly affected the stock market conditions [21]. So COVID-19, which is more widespread than SARS, will have an impact on stock market conditions.

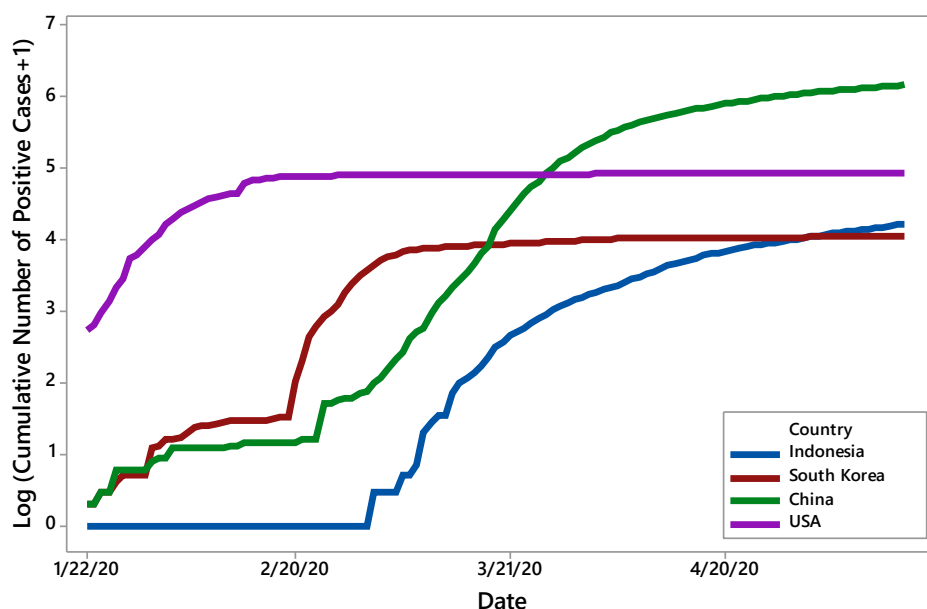


Figure 2. Time series plots the cumulative number of COVID-19 cases in Indonesia, South Korea, China and the United States

The first case of COVID-19 in each country is different. Table 2 shows the first cases of COVID-19 in the United States, Indonesia, China, and South Korea based on WHO and John Hopkins University data. After the first COVID-19 case will sooner or later have an impact on stock prices in the world and the country concerned. Figure 3 shows a time series plot of composite stock prices in 3 countries in the world, namely the United States, China, and South Korea.

Table 2. The first COVID-19 case in each country

Country	First Case
USA	January 21, 2020
South Korea	January 20, 2020
China	December 27, 2019
Indonesia	March 2, 2020

Based on Figure 3, it can be seen in three major countries, namely South Korea, the United States, and China after the presence of COVID-19 has a significant impact on composite stock prices in each country. The effect of COVID-19 on stock prices is not felt directly in China. The decline in composite stock prices in China after the lowest COVID-19 at around early February. At that time was the peak period of the spread of COVID-19 in China, the average increase in the number of daily cases in early February was quite high. The first case of COVID-19 also indirectly affected the United States and South Korea. Before the COVID-19, composite stock prices tended to be stable. After the COVID-19 experienced a decline and the lowest decline occurred at the end of March 2020. The decline in the stock price at the end of March 2020 was also experienced by all countries including China. In the period around the end of March, the spread of the COVID-19 virus was quite massive, especially in developed countries including the United States and South Korea. As a result, it has an impact on the economic conditions of all countries in the world. But from April to May 2020

conditions in the economy are better, it can be seen from the composite stock price that has begun to rise.

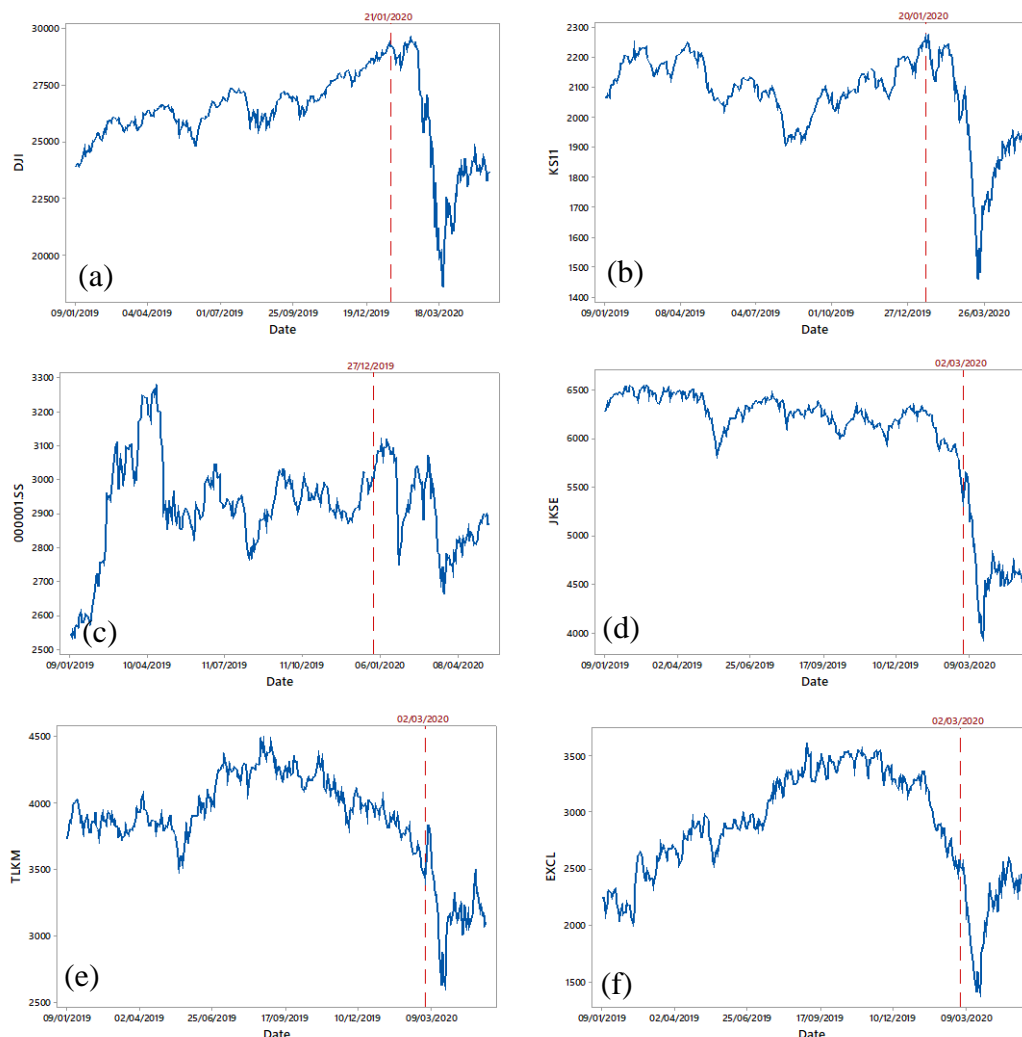


Figure 3. Time series plot of composite stock prices DJI (a), KS11 (b), 000001.SS (c), JKSE (d), TLKM (e), and EXCL (f)

The existence of COVID-19 also has an impact on the Indonesian economy. Indonesia's first positive case on March 2, 2020. After this first case, the composite stock price of Indonesian and individual stock price, especially the telecommunications sector, i.e. Telkom and XL, tended to fall and the lowest at the end of March. This condition is the same as that experienced by other countries. After that, from the beginning of April until now in mid-May 2020, Indonesia's composite stock index tends to rise, although the increase is not significant. Something different is precisely in the individual stock price index in the telecommunications sector, i.e. Telkom and XL. After dropping at the end of March, the prices of Telkom and XL shares have risen even though the condition is now almost the same as before the COVID-19 in Indonesia. To find out the effect of COVID-19 on composite and individual stock prices in the world and Indonesia, an intervention analysis will be conducted.

B. Intervention Analysis on Stock Prices Data

In analyzing the intervention consists of several stages. Intervention analysis begins with ARIMA modeling for data before the intervention. After that, the intervention analysis modeling is done by identifying the order b, r, and s, then estimating parameters, checking diagnoses, and forecasting. In this study, an analysis will be conducted on the effect of COVID-19 intervention on stock prices in the world and Indonesia.

Before an intervention analysis is performed, the ARIMA model is first searched for data before the intervention. The intervention began when there was the first case of COVID-19 in the country. ARIMA modeling begins with the identification of data stationarity. The combined share price of JKSE (Indonesia) is not stationary in the mean, so there is a differencing lag 1. After differencing the first lag obtained, the data is stationary. The ARIMA model before intervention (the first COVID-19 case in Indonesia) for the JKSE stock price is ARIMA (0,1,1). The parameters in the model have been significant. The residuals of the ARIMA model for JKSE stock prices have met the assumptions of white noise and are normally distributed.

Using the same procedure as ARIMA modeling on JKSE stock prices with box-Jenkins, ARIMA models are obtained for composite stock price data in other countries, i.e. South Korea, China, the United States, and individual stock prices

in Indonesia, i.e. Telkom (TLKM) and XL (EXCL) before COVID-19 cases appeared in each country. Based on Table 3 for all stock prices, namely KS11, 000001.SS, DJI, TLKM, and EXCL, the ARIMA models obtained are the same, namely the ARIMA model (0,1,0). The model is a random walk model that is the current observation influenced by previous observations. It can be seen that the residuals have met the white noise assumption. Whereas the normal distribution assumption has not been fulfilled. One cause of the assumption that residuals are not normally distributed is the existence of outliers in the data.

Table 3. ARIMA models at each stock price

Stock Prices	ARIMA Model	White Noise Assumption	Normal Distribution Assumption			
KS11	ARIMA(0,1,0)	Yes	No			
000001.SS	ARIMA(0,1,0)	Yes	No			
DJI	ARIMA(0,1,0)	Yes	No			
JKSE	ARIMA(0,1,1)	Yes	Yes			
TLKM	ARIMA(0,1,0)	Yes </tr <tr> <td>EXCL</td> <td>ARIMA(0,1,0)</td> <td>Yes</td> <td>No</td> </tr>	EXCL	ARIMA(0,1,0)	Yes	No
EXCL	ARIMA(0,1,0)	Yes	No			

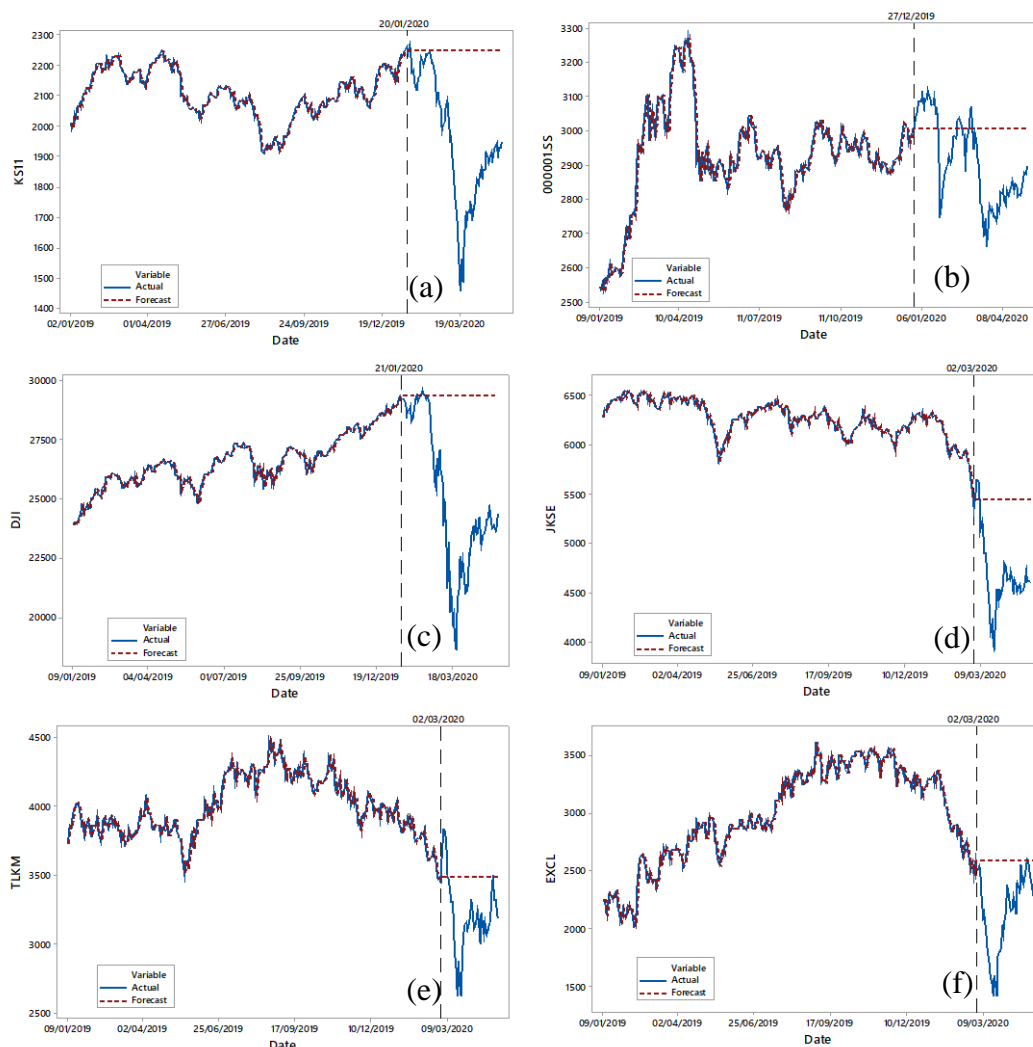


Figure 4. Comparison of actual data and forecast results using ARIMA before intervention on stock prices KS11 (a), 000001.SS (b), DJI (c), JKSE (d), TLKM (e) and EXCL (f) in the in-sample data

From Figure 4, it can be seen that the general stock price for forecast results using the ARIMA method before the first COVID-19 case in each country follows the actual data. But after the first COVID-19 case in each country, the forecast results did not follow the actual data pattern. It indicates the influence/intervention of COVID-19 on stock prices both in the world and in Indonesia. To find out the effect of the intervention a standardized residual/response function plot is obtained to get the order b, r, s from the intervention model.

After the forecast results are obtained using the ARIMA model, it can be calculated the standardized residual value based on the ARIMA model for each stock price data. It was done to identify the intervention model. In intervention modeling, the value of the intervention variable for each country is different. It is because the first case of COVID-19 is

different in each country. The intervention model used is a step function, because the COVID-19 case is still running. The value of this intervention variable will be one if $t \geq T$ and zero if $t < T$. T is an exchange day where there is the first case of COVID. The values can be seen in Table 4.

Table 4. The value of the intervention variable in each stock price data

Stock Price	Equal 1	Equal 0
KS11 (X_1)	$t \geq 259$	$t < 259$
000001.SS (X_2)	$t \geq 235$	$t < 235$
DJI (X_3)	$t \geq 260$	$t < 260$
JKSE (X_4)	$t \geq 295$	$t < 295$
TLKM (X_5)	$t \geq 295$	$t < 295$
EXCL (X_6)	$t \geq 295$	$t < 295$

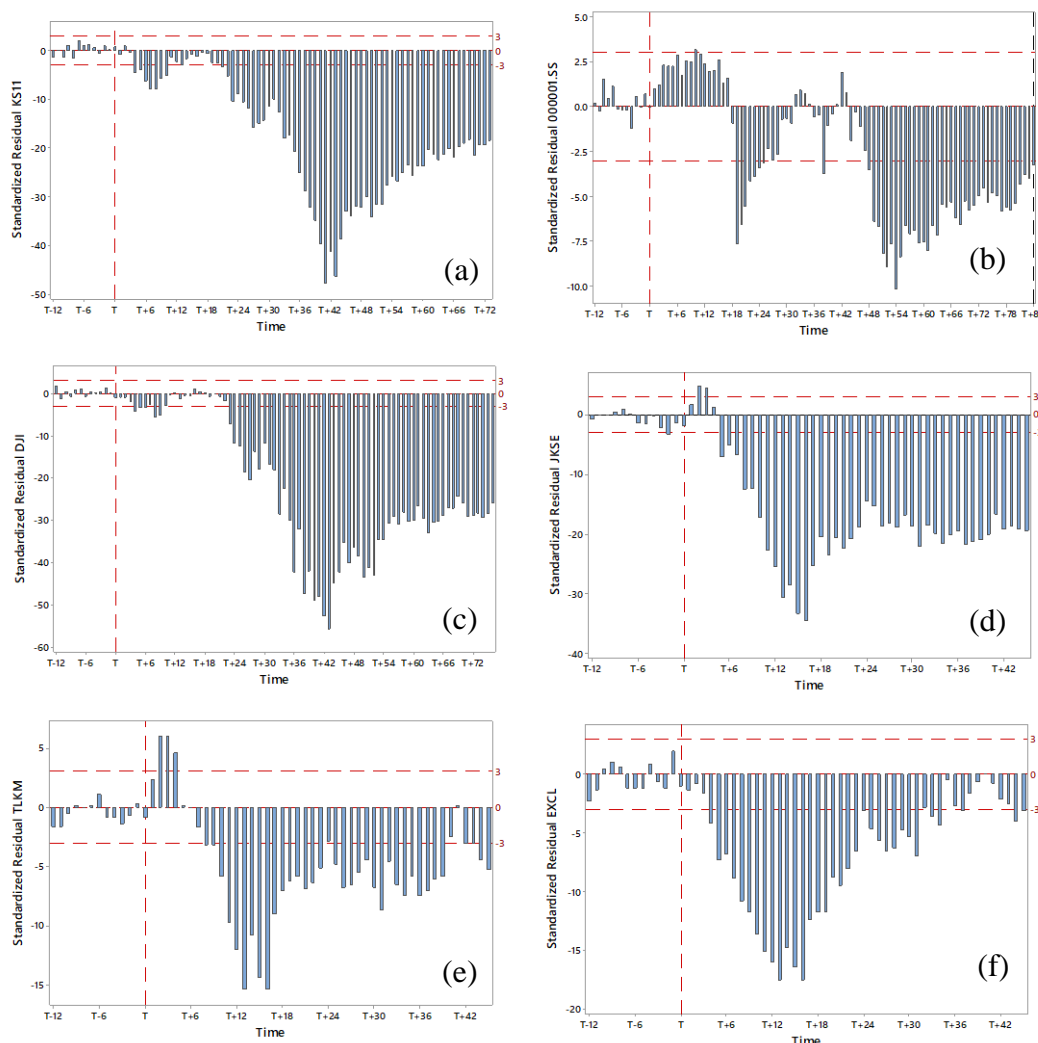


Figure 5. Standardized residual/response function plot at stock prices KS11 (a), 000001.SS (b), DJI (c), JKSE (d), TLKM (e), and EXCL (f)

From Figure 5, it can be seen that the presence of COVID-19 intervention when T causes an effect in stock prices. In general, the effects are not immediate, but some time after the first case of COVID-19 in the country. Based on Figure 5, it can be obtained order b , r , s , p , d , and q for the intervention model of each stock price both composite stock and individual stock. Table 5 shows the order of the intervention model for each stock price based on the standardized residual/response function plot in Figure 5. The model obtained meets the white noise assumption, but the assumption of a normal distribution is not fulfilled. It is due to an outlier in the data. All parameters in the model have been significant.

In general, from the results of the intervention analysis in Table 5, the composite stock price in various countries is greatly influenced by the presence of COVID-19. Based on Table 6, it can be seen that the presence of COVID-19 hurts composite stock prices in almost all countries, both those that have passed the peak periods, such as China and South Korea, and those that are still in peak periods, such as Indonesia and the United States. Negative impacts are not felt directly after the first case of COVID-19 in each country, and the negative impacts caused are permanent. From Figure 6

and Table 6, it can be seen that simultaneously the highest decline in each stock price occurred at the end of March 2020, namely between 18 March 2020 and 24 March 2020. It is related to global economic conditions that were shaken by the presence of COVID-19. Around the end of March 2020, the spread of COVID-19 was quite widespread in almost all countries. After the end of March, the decline that began to decrease is due to the economic stimulus provided by each country. It can be seen from Figure 7 that the composite stock price in Indonesia experienced the largest decrease in percentage compared to the composite stock price in South Korea, China, and the United States.

Table 5. Order of the intervention model at each stock price

Stock Price	b	r	s	p	d	q
KS11	4	0	[7, 19, 23, 29, 32, 36, 37, 38, 39, 40, 41, 46, 49, 57]	0	1	0
000001.SS	19	0	[10, 13, 19, 20, 23, 25, 26, 30, 35, 37]	0	1	0
DJI	4	0	[4, 6, 19, 24, 26, 27, 29, 30, 31, 32, 33, 34, 35, 36, 40, 42, 51]	0	1	[3, 4, 5, 6, 7, 8]
JKSE	2	0	[2, 3, 6, 8, 9, 11, 13, 15, 16, 22, 24, 30, 39]	0	1	0
TLKM	2	0	[3, 8, 9, 11, 12, 13, 15, 28, 38, 39, 40]	0	1	0
EXCL	4	0	[1, 3, 4, 6, 11, 13, 16, 20, 27, 28, 31, 32]	0	1	0

Table 6. Comparison of the impact of COVID-19 on each stock price

Stock Price	The First COVID-19 Case in Each Country	First Impact Felt	Lowest Stock Prices Decline	Impact Type
KS11	20 January 2020	After 4 exchange days	After 41 exchange days (19 March 2020)	Negative and Permanent
000001.SS	27 December 2019	After 19 exchange days	After 54 until 55 exchange days (23-24 March 2020)	Negative and Permanent
DJI	21 January 2020	After 4 exchange days	After 40 until 43 exchange days (18-23 March 2020)	Negative and Permanent
JKSE	2 March 2020	After 2 exchange days	After 15 until 16 exchange days (23-24 March 2020)	Negative and Permanent
TLKM	2 March 2020	After 2 exchange days	After 13 exchange days (19 March 2020)	Positive up to 4 exchange days, negative from 5 exchange days to 40 exchange days, positive on 41 exchange days and negative from 42 exchange days to 45 exchange days
EXCL	2 March 2020	After 4 exchange days	After 16 exchange days (24 March 2020)	Negatives up to 31 exchange days and positive from 32 exchange days to 45 exchange days

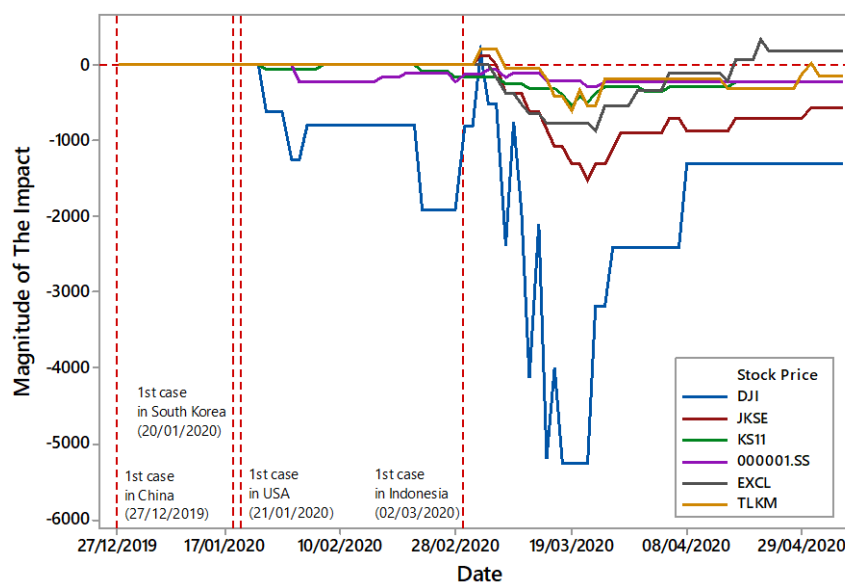


Figure 6. Plot of the magnitude of the impact of COVID-19 on each stock price

The existence of COVID-19 also affects individual stock prices in Indonesia including the telecommunications sector

such as TLKM and EXCL. From Figure 6, it can be seen that this sector is quite affected and has decreased until the end of March 2020. Around the end of April 2020, individual stock prices showed a fairly positive trend. Even though the condition of Indonesia's composite stock prices is still experiencing a decline. It is because during the COVID-19 outbreak all sectors using an online system that resulted in the telecommunications sector's stock price becoming stronger. Table 6 shows in detail the impact of COVID-19 on each stock price.

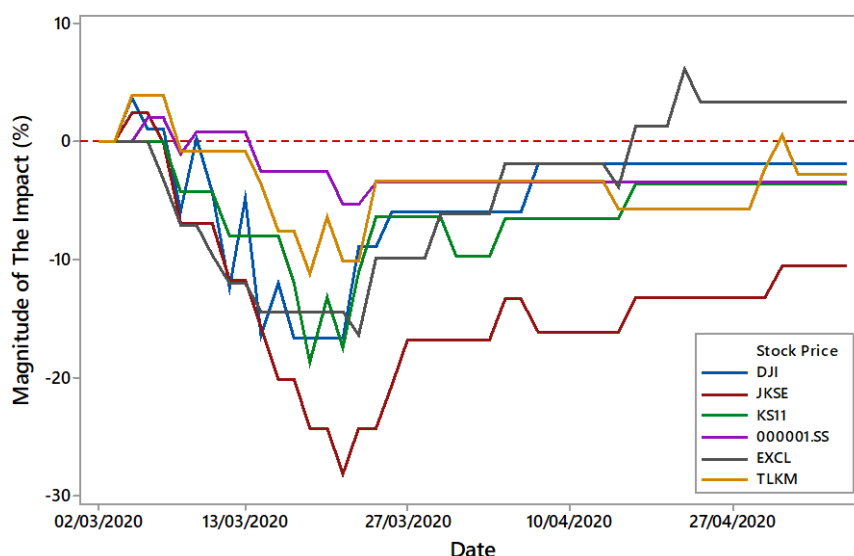


Figure 7. Plot of the magnitude of the impact of COVID-19 on each stock price after the first case in Indonesia (March 2, 2020) in %

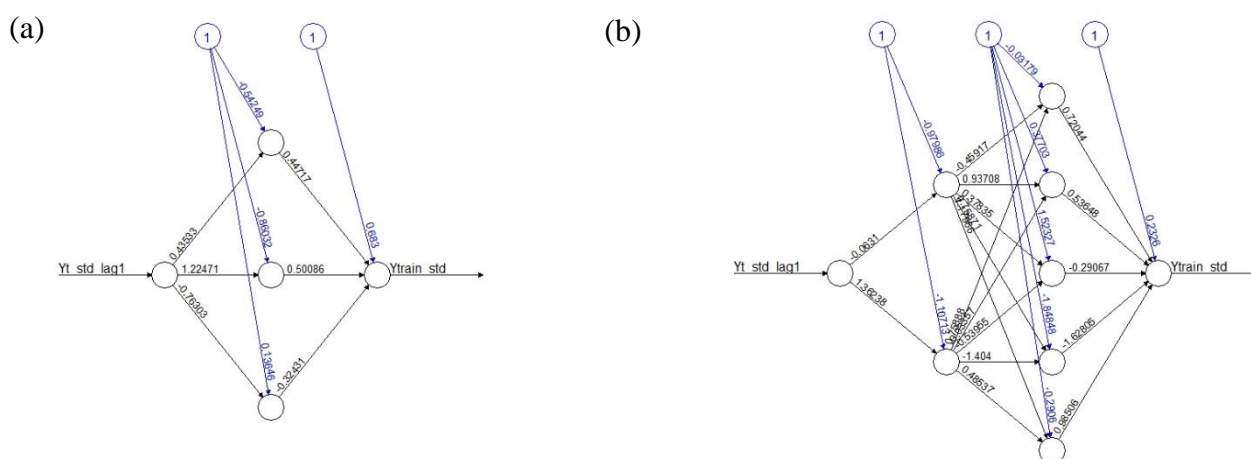


Figure 8. NN architecture of DJI stock price with input lag 1 (a) and DLNN architecture of DJI stock price with input lag 1 (b)

Table 7. The optimum number of neurons in the NN and DLNN models

Stock Price	NN (Input Lag 1)	NN (Input Lag 1 and Dummy)	DLNN (Input Lag 1)	DLNN (Input Lag 1 and Dummy)
KS11	1	4	2, 5	10, 1
000001.SS	3	5	4, 3	3, 4
DJI	3	2	2, 5	1, 10
JKSE	3	1	4, 3	1, 10
TLKM	4	1	10, 1	2, 5
EXCL	3	3	10, 1	10, 1

C. Machine Learning Model on Stock Prices Data

Besides using statistical methods in forecasting stock prices, machine learning models are also used, namely neural network (NN) and deep learning neural network (DLNN). Inputs used in the NN and DLNN models are two scenarios namely, first lag and first lag with the dummy variable from the intervention (the COVID-19 case). Before the data is used in the NN and DLNN models, preprocessing is done first by changing the data to between 0 and 1. The tanh activation function is used in the NN and DLNN models. The NN model uses one hidden layer and is tried 1, 2, 3, 4, 5,

and 10 neurons in the hidden layer. Whereas, in the DLNN model using two hidden layers, a combination of neurons 1, 2, 3, 4, 5, and 10 is tried in each hidden layer. Table 7 shows the optimum number of neurons in the NN and DLNN models. The NN and DLNN architectures for each stock price are different. Figure 8 shows the architecture of the NN and DLNN models of the DJI stock price with input first lag, while Figure 9 shows the architecture of the NN and DLNN model of the DJI stock price with input lag 1 and the dummy variable of COVID-19.

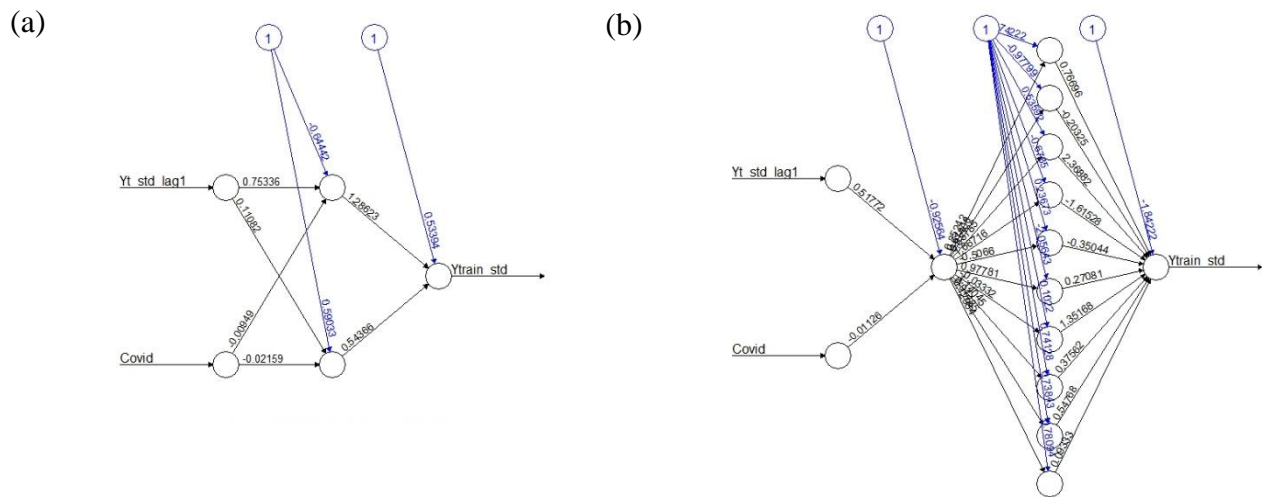


Figure 9. NN architecture of DJI stock price with input lag 1 and dummy variable of COVID-19 (a) and DLNN architecture of DJI stock price with input lag 1 and dummy variable of COVID-19 (b)

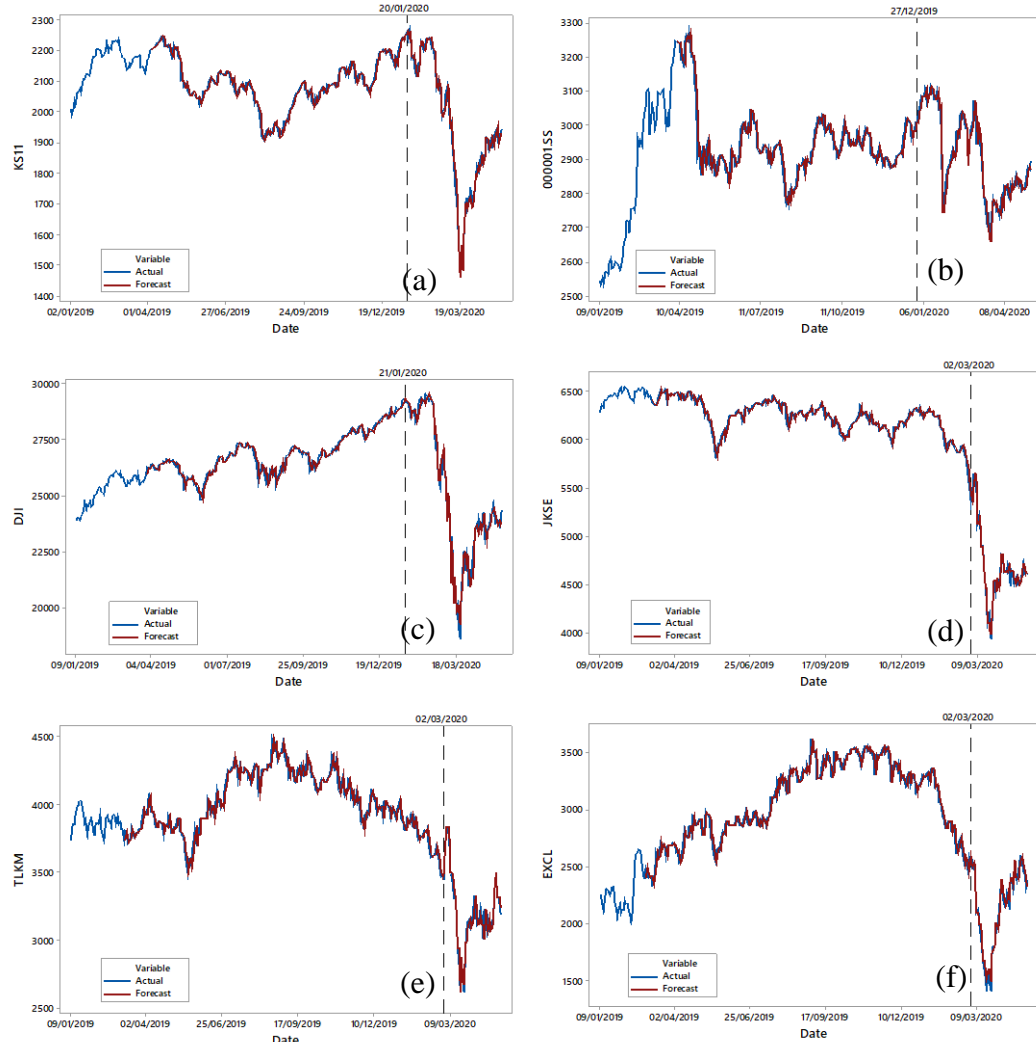


Figure 10. Comparison of actual data and forecast results using intervention analysis on stock prices KS11 (a), 000001.SS (b), DJI (c), JKSE (d), TLKM (e) and EXCL (f) in the in-sample data

Table 8. Comparison of Intervention Analysis and Machine Learning models

Stock Prices	Data	RMSE				
		Intervention Model	NN (Input Lag 1)	DLNN (Input Lag 1)	NN (Input Lag 1 and Dummy)	DLNN (Input Lag 1 and Dummy)
KS11	IS	18.96	26.60	25.66	24.62	24.52
	OS1	19.14	18.34*	20.18	25.01	20.96
	OSk	24.19	29.31	29.83	27.27	27.06
000001.SS	IS	30.15	37.01	36.60	36.85	36.53
	OS1	14.27	14.17	14.23	13.75	13.73*
	OSk	15.45	21.01	34.86	15.89	11.31
DJI	IS	265.66	432.18	432.02	424.54	424.03
	OS1	501.50	457.27*	457.67	483.77	484.32
	OSk	526.42	544.88	535.83	681.91	713.29
JKSE	IS	49.87	63.81	63.49	64.39	64.34
	OS1	32.72	30.91	29.81	29.52*	29.67
	OSk	62.46	44.46	42.09	32.63	33.54
TLKM	IS	58.99	70.38	70.42	70.34	65.19
	OS1	70.62	69.52	64.05*	64.48	66.01
	OSk	64.51	86.99	72.38	64.54	62.65
EXCL	IS	64.97	77.14	77.17	76.65	72.95
	OS1	78.26	77.37	74.64	79.08	73.72*
	OSk	73.74	89.25	99.65	151.97	106.32

IS=In-sample, OS1=Out-of-Sample (1-Step), OSk=Out-of-Sample (k-Step), *=Smallest RMSE

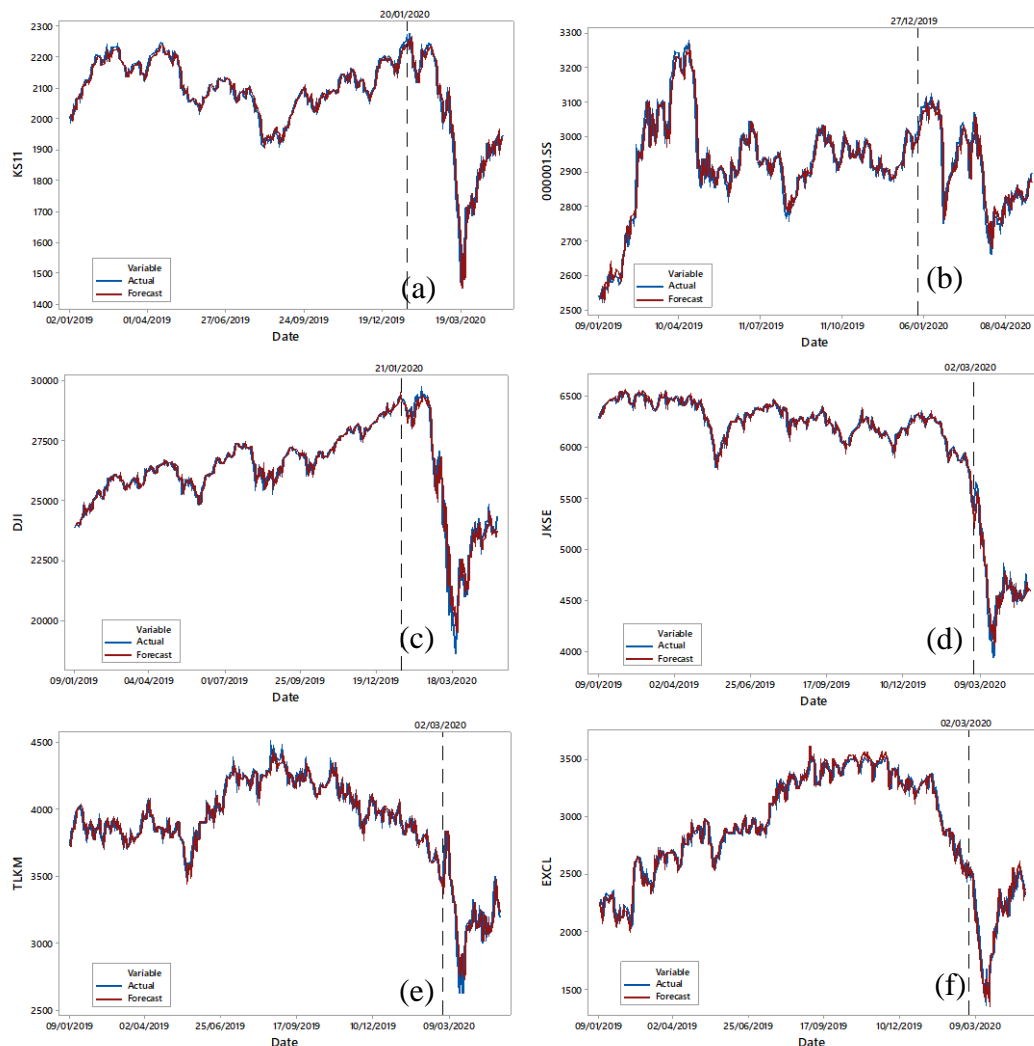


Figure 11. Comparison of actual data and forecast results using machine learning (NN and DLNN) on stock prices KS11 (a), 000001.SS (b), DJI (c), JKSE (d), TLKM (e) and EXCL (f) in the in-sample data

D. Forecasting Stock Prices

Figure 10 shows the comparison of actual data and forecasts on in-sample data using the intervention model for each stock price. It can be seen that the forecast results with the intervention model have followed the actual data. The results of this forecast are better than the results of the forecast in Figure 4. Then a comparison is made between the intervention model and the machine learning model, i.e. NN and DLNN. Figure 11 shows the comparison of actual data and forecasts on in-sample data using machine learning methods for each stock price. The machine learning method used is the best NN and DLNN method based on RMSE values.

In out-of-sample data, forecasting is done with a 1-step and k-step scheme both in the intervention analysis and machine learning model. 1-step forecasting is done every one period ahead, while the k-step forecasting is directly carried out as many as the k-step ahead. Figure 12 shows the comparison of actual data and forecasts on out-of-sample data using the intervention model for each stock price. It can be seen that forecasts with k-step schemes tend to produce constant forecast results, while predictions with 1-step schemes tend to produce forecast results that follow actual data. The same condition also occurs in the forecast results of machine learning methods, it can be seen in Figure 13.

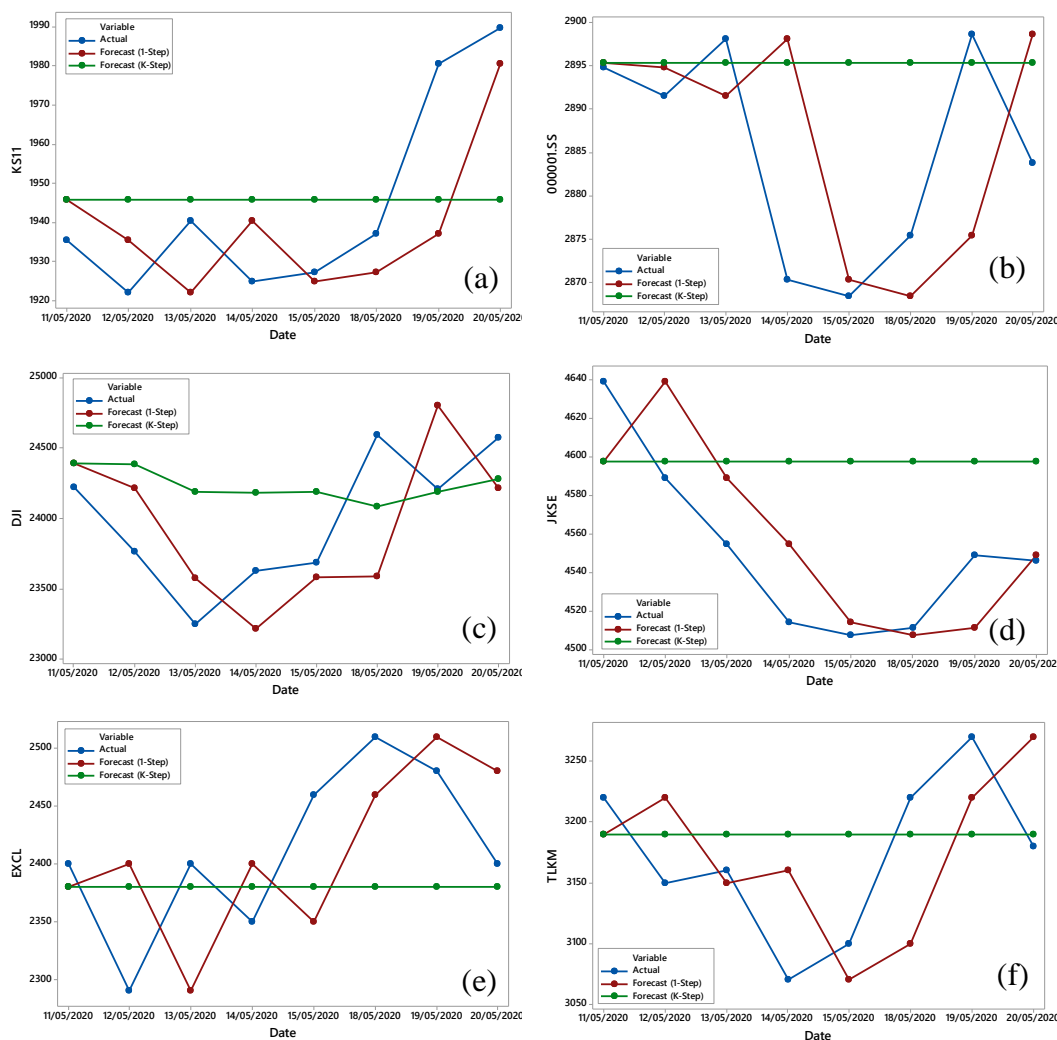


Figure 12. Comparison of actual data and forecast results using intervention analysis on stock prices KS11 (a), 000001.SS (b), DJI (c), JKSE (d), EXCL (e) and TLKM (f) in the out-of-sample data

To evaluate the model, RMSE is used in the out-of-sample data with 1-step ahead forecasting. Based on Table 8, it can be seen that for all stock price data, i.e. KS11, 000001.SS, DJI, JKSE, TLKM, and EXCL, machine learning models namely NN and DLNN are better than the intervention model. But the difference in the RMSE value is not large. If forecasting using the k-step scheme on out-of-sample data is known that the intervention model is better for three stock prices, and the machine learning model is better for three stock prices. In Table 8, based on the RMSE value in the in-sample data for all stock price data, the intervention model is better than the machine learning model.

IV. DISCUSSION

Based on the results of the intervention analysis, the composite stock price in various countries such as South Korea, China, the USA, and Indonesia is greatly influenced by the presence of COVID-19. These results are in line with studies

conducted by [22] and [23] using different methods. In general, COVID-19 hurts the price of both composite and individual stock prices. Simultaneously the highest decline in each stock price occurred at the end of March 2020, between 18 March 2020 and 24 March 2020. This simultaneous decline was caused by several things such as, global economic conditions that were shaken by the presence of COVID-19 and the widespread COVID-19 in the whole world. The individual stock prices of the telecommunications sector in Indonesia, such as TLKM and EXCL showed a fairly positive trend at the end of April 2020. The first case of COVID-19 in each country did not have a direct impact on all share prices.

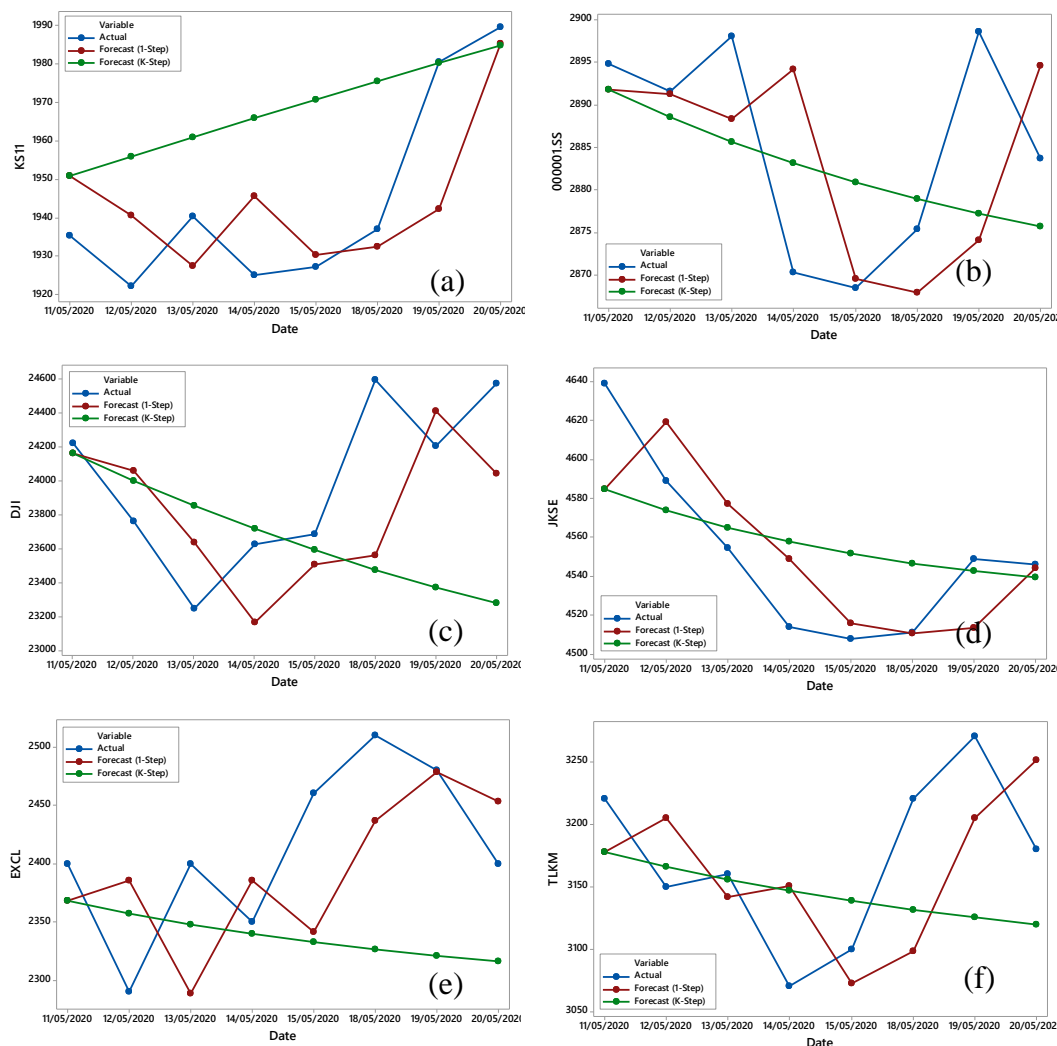


Figure 13. Comparison of actual data and forecast results using machine learning (NN and DLNN) on stock prices KS11 (a), 000001.SS (b), DJI (c), JKSE (d), EXCL (e) and TLKM (f) in the out-of-sample data

The RMSE value in the in-sample data shows that the intervention analysis is better than the machine learning model (NN and DLNN) for all stock prices. While the forecast results with a 1-step scheme in out-of-sample data show that the machine learning model is better than the intervention analysis for all stock prices. These results are in line with research conducted by [24], [25], [26], and [27] that showed machine learning models are better than the classical time series models. In general, the intervention model is better for an explanation, especially knowing the impact of a phenomenon [13], while machine learning is better for forecasting. Results of predictions on out-of-sample data with intervention analysis and machine learning show that forecasts with k-step schemes tend to be constant and do not follow actual data. While the forecast results on out-of-sample data with a 1-step scheme show the forecast results with the intervention model and machine learning following the actual data. It shows that stock price forecasting is better for short-term forecasts, i.e. 1-step ahead compared to long-term forecasts, i.e. k-step ahead.

V. CONCLUSION

In this study, an analysis was conducted to determine the impact of COVID-19 on stock market conditions in the world and Indonesia. In general, the COVID-19 outbreak hurt stock prices. The impact is not directly after the first case of COVID-19 in each country. Based on the modeling of the intervention analysis until the latest data in this analysis, i.e. May 8, 2020, the COVID-19 resulted in a decrease in the stock prices. The decrease happened, both in countries that have passed peaks, such as China (000001.SS) and South Korea (KS11), and those still in peak periods such as the United States (DJI) and Indonesia (JKSE). In general, COVID-19 resulted in a decline in stock prices from the beginning of March 2020

to the lowest decline at the end of March 2020, namely between March 18 and March 24, 2020. This decline in stock prices also had an impact on the stock prices of individual telecommunications sectors such as TLKM and EXCL. This sector tends to show a positive trend after a significant decline at the end of April 2020. For all stock prices forecast by the machine learning model (NN and DLNN) have better accuracy than the intervention model in out-of-sample data. While the forecast results with the intervention model have better accuracy compared to the machine learning model on in-sample data. For further research, it can be used multiple inputs in the intervention analysis models, i.e. the addition of lockdown implementation variables. Besides, statistical models and other machine learning models and hybrid models can also be used to forecast stock prices.

REFERENCES

- [1] T. Zach, "Political Events and The Stock Market: Evidence from Israel," *International Journal of Business*, vol. 8, no. 3, pp. 10.2139/ssrn.420242, 2003.
- [2] O. Kowalewski and P. Spiewanowski, "Stock Market Response to Potash Mine Disasters," *Journal of Comodity Market*, pp. 100124, 2020.
- [3] A. Bash and K. Alsaifi, "Fear from uncertainty: An Event Study of Khashoggi and Stock Market Returns," *Journal of Behavioral and Experimental Finance*, vol. 23, pp. 54-58, 2019.
- [4] M. Guo, Y. Kuai and X. Liu, "Stock Market Response to Environmental Policies: Evidence from Heavily Polluting Firms in China," *Economic Modelling*, vol. 86, pp. 306-316, 2020.
- [5] G. Wong, "Has SARS Infected The Property Market? Evidence from Hong Kong," *Journal of Urban Economics*, vol. 63, no. 1, pp. 74-95, 2008.
- [6] C. D. Chen, C. C. Chen, W. W. Tang and B. Y. Huang, "The Positive and Negative Impacts of The SARS Outbreak: A Case of the Taiwan Industries," *The Journal of Developing Areas*, vol. 43, no. 1, pp. 281-293, 2009.
- [7] B. Buetre, Y. Kim, Q. T. Tran and D. Gunasekera, "Avian Influenza: Potential Economic Impact of a Pandemic on Australia," *Australian Commodities: Forecasts and Issues*, vol. 13, no. 2, pp. 351-359, 2006.
- [8] R. Ichev and M. Marinic, "Stock prices and geographic proximity of information: Evidence from the Ebola outbreak," *International Review of Financial Analysis*, vol. 56, pp. 153-166, 2018.
- [9] M. P. Wilson and A. S. Jack, "Coronavirus Disease 2019 (COVID-19) in Neurology and Neurosurgery: A Scoping Review of The Early Literature," *Clinical Neurology and Neurosurgery*, vol. 193, pp. 105866, 2020.
- [10] N. A. Sansa, "The Impact of The COVID-19 on The Financial Markets: Evidence from China and USA," *Electronic Research Journal of Social Sciences and Humanities*, vol. 2, pp. 29-39, 2020.
- [11] A. M. Al-Awadhi, K. Al-Saifi, A. Al-Awadhi and S. Alhamadi, "Death and Contagious Infectious Diseases: Impact of the COVID-19 Virus on Stock Market Returns," *Journal of Behavioral and Experimental Finance*, vol. 27, pp. 100326, 2020.
- [12] W. W. Wei, *Time Series Analysis Univariate and Multivariate Methods* (2nd ed), United States of America: Pearson Education, Inc, 2006.
- [13] G. E. P. Box and G. C. Tiao, "Intervention Analysis With Applications to Economic and Environmental Problems," *Journal of the American Statistical Association*, vol. 70, pp. 70-79, 1975.
- [14] S. Vujic, J. J. F. Commandeur and S. J. Koopman, "Intervention Time Series Analysis of Crime Rates: The Case of Sentence Reform in Virginia," *Economic Modelling*, vol. 57, pp. 311-323, 2016.
- [15] G. E. P. Box, G. M. Jenkins, G. C. Reinsel and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, Hoboken: Wiley, 2015.
- [16] B. L. Bowerman and O'Connel, *Forecasting and Time Series: An Applied Approach* 3rd edition, California: Duxbury Press, 1993.
- [17] X. Liu, H. Zhang, X. Kong and K. Y. Lee, "Wind Speed Forecasting Using Deep Neural Network with Feature Selection," *Neurocomputing*, vol. 397, pp. 393-403, 2020.
- [18] Z. Yang, M. Mourshed, K. Liu, X. Xu and S. Feng, "A Novel Competitive Swarm Optimized RBF Neural Network Model for Short-Term Solar Power Generation Forecasting," *Neurocomputing*, vol. 397, pp. 415-421, 2020.
- [19] J. Faraway and C. Chatfield, "Time Series Forecasting with Neural Network: A Comparative Study Using The Airline Data," *Applied Statistics*, vol. 47, pp. 231-250, 1998.
- [20] E. Chong and S. H. Zak, *An Introduction to Optimize*, Canada: John Wiley & Son, 2001.
- [21] M. H. Chen, S. S. Jang and W. G. Kim, "The Impact of the SARS Outbreak on Taiwanese Hotel Stock Performance: An Event-Study Approach," *International Journal of Hospitality Management*, vol. 26, no. 1, pp. 200-212, 2007.
- [22] H. Y. Liu, A. Manzoor, C. Y. Wang, L. Zhang and Z. Manzoor, "The COVID-19 Outbreak and A Stock Markets Response," *International Journal of Environmental Research and Public Health*, vol. 17, no. 8, pp. 2800, 2020.
- [23] M. Ali, N. Alam and S. A. R. Rizvi, "Coronavirus (COVID-19) — An Epidemic or Pandemic for Financial Markets," *Journal of Behavioral and Experimental Finance*, vol. 27, pp. 100341, 2020.
- [24] S. L. Ho, M. Xie and T. N. Goh, "A Comparative Study of Neural Network and Box-Jenkins ARIMA modelling in Time Series Prediction," *Computer & Industrial Engineering*, vol. 42, pp. 371-375, 2002.
- [25] B. Fanoodi, B. Malmir and F. F. Jahantigh, "Reducing Demand Uncertainty in The Platelet Supply Chain Through Artificial Neural Networks and ARIMA Models," *Computers in Biology and Medicine*, vol. 113, pp. 103415, 2019.

- [26] M. Matyjaszek, P. R. Fernandez, A. Krzemien, K. Wodarski and G. F. Valverde, "Forecasting Coking Coal Prices by Means of ARIMA Models and Neural Networks, Considering The Transgenic Time Series Theory," *Resources Policy*, vol. 61, pp. 283-292, 2019.
- [27] F. S. Lasheras, F. J. C. Juez, A. S. Sanchez, A. Krzemien and P. S. Fernandez, "Forecasting The COMEX Copper Spot Price by Means of Neural Network and ARIMA Models," *Resources Policy*, vol. 45, pp. 37-43, 2015.