© Universiti Tun Hussein Onn Malaysia Publisher's Office

IJIE

The International Journal of Integrated Engineering

http://penerbit.uthm.edu.my/ojs/index.php/ijie

ISSN : 2229-838X e-ISSN : 2600-7916

COVID-19 Confirmed Cases Forecasting in Malaysia Using Linear Regression and Holt's Winter Algorithm

Hudzaifah Hasri[1], Siti Armiza Mohd Aris[1]*, Robiah Ahmad[1], Celia Shahnaz[2]

[1]Razak Faculty of Technology and Informatics,
Universiti Teknologi Malaysia, Kuala Lumpur, 54100, MALAYSIA

[2]Electrical and Computer Engineering,
Bangladesh University of Engineering and Technology, Dhaka, 1205, BANGLADESH

*Corresponding Author

DOI: https://doi.org/10.30880/ijie.2023.15.03.006
Received 31 October 2022; Accepted 29 December 2022; Available online 31 July 2023

Abstract: The 2019 coronavirus disease pandemic (COVID-19) has emerged and is spreading rapidly over the world. Therefore, it may be highly significant to have the general population tested for COVID-19. There has been a rapid surge in the use of machine learning to combat COVID-19 in the past few years, owing to its ability to scale up quickly, its higher processing power, and the fact that it is more trustworthy than people in certain medical tasks. In this study, we compared between two different models: the Holt's Winter (HW) model and the Linear Regression (LR) model. To obtain the data set of COVID-19, we accessed the website of the Malaysian Ministry of Health. From January 24th, 2020, through July 31st, 2021, daily confirmed instances were documented and saved in Microsoft Excel. Case forecasts for the next 14 days were generated in the Waikato Environment for Knowledge Analysis (WEKA), and the accuracy of the forecasting models was measured by means of the Mean Absolute Percentage Error (MAPE). According to the lowest value of performance indicators, the best model is picked. The results of the comparison demonstrate that Holt's Winter showed better forecasting outcome than the Linear Regression model. The obtained result depicted the forecasted model can be further analyzed for the purpose of COVID-19 preparation and control.

Keywords: Linear regression, Holt's Winter, prediction, COVID-19, Malaysia, WEKA

1. Introduction

Recently the world has been experiencing an outbreak of the dangerous coronavirus disease 2019 (COVID-19). This novel virus, designated COVID-19, is deadly. Most infected persons will have short-term respiratory symptoms that are easily treatable with rest and treatment, and the virus will eventually leave their systems on its own. Some, though, will become gravely ill and necessitate professional medical care. According to the World Health Organization (WHO), China's Hubei Province, where Wuhan is located, has seen an increase in pneumonia with no clear explanation [1]. From there on, COVID-19 has become increasingly widespread, posing a threat to global public health. More than 118,000 cases have been reported in 114 countries, and 4,292 people have died since March 11th, 2020 [2]. Aside from China, the European region is also critically affected by the virus, especially in Italy and France, with 10,149 and 1,174 cases, respectively [2].

*Corresponding author: armiza.kl@utm.my

Subsequently, the WHO announced COVID-19 as a global pandemic following an outbreak assessment based on these numbers [2].

COVID-19 was first discovered in Malaysia on January 25^{th}, 2020 [3]. The total number of confirmed COVID-19 cases began to rise over time. As of May 11^{th}, 2021, the Ministry of Health in Malaysia (MOH) reports 448,457 instances of COVID-19, including 3,973 new cases [4]. At the end of March 2020, Malaysia's population of people infected with COVID-19 disease rose to thousands. Between June 2020 and August 2020, Malaysia reached the lowest number of cases of COVID-19. Sadly, the number rise resumed in September 2020 and has continued. Since the discovery of COVID-19, scientists from various disciplines have studied this novel virus. Public health decision-making and resource allocation can benefit significantly from COVID-19 case forecasting in their efforts to lessen the pandemic's impact on morbidity and mortality [5].

According to [6], artificial intelligence (AI) and machine learning (ML) are two promising technologies that a wide variety of healthcare providers can employ. These technologies can scale up competently, handle data more quickly, and be more reliable than humans in specific healthcare operations. As a result, healthcare providers worldwide deployed a wide range of ML and AI solutions to predict and fight the COVID-19 pandemic and overcome the challenges it caused to humankind. A study by A. Sözen, A. D. Tuncer, and F. Kazancoglu indicated the confirmed cases of COVID-19 using Long-Short-Term Memory (LSTM) and Nonlinear Autoregression Neural Network (NARNN) approaches [6]. In the following section, Linear Regression (LR), Multilayer Perceptron (MLP), and Vector Autoregression (VAR) were examined as potential methods for predicting confirmation, dead, and recovered patients [7]. According to the findings, the MLP method outperforms the LR and VAR methods in terms of accuracy. N. Alballa and I. Al-Turaiki had studied that for COVID-19 diagnosis and prognosis, the LR model is the most commonly used algorithm [8]. There are many reasons for this, and the most important one is that the LR model is simple to calculate and can handle continuous numerical data.

MLP and an ANFIS (adaptive network-based fuzzy inference system) were used by Ardabili et al. to make predictions about confirmed COVID-19 cases [9]. Without relying on the probabilities often related to epidemiology models, the models showed encouraging results in time series prediction. The total rise in infection cases was predicted using both XGBoost and MultiOutputRegressor in a study by Y Suzuki et. al [10]. The model's accuracy rate was 82.4 percent, which is relatively high. To estimate the total number of confirmed patients who will be susceptible to the disease in the future globally, E. Gothai et al. [11] trained three distinct ML algorithms. Linear Regression (LR), Support Vector Machine (SVM) and Holt's Winter (HW) are the algorithms used. For predicting future global confirmed cases with an accuracy of 87%, the author found HW model outperformed LR and SVM. The sudden increase in COVID-19 cases has put a strain on healthcare facilities around the world. Detecting COVID-19 in the general population ahead of time would be very useful, especially for predicting how many patients would contract the virus during the anticipated peak transmission period.

Predictions can be made in numerous ways, one of which is the HW method for predicting time series data. A researcher in India used HW and Autoregressive Integrated Moving Average (ARIMA) models to forecast COVID-19 cases [12]. The study resulting in both models gave good accuracy and can be utilized to predict future cases. I Djakaria and S E Saleh came out with a HW model with smoothing parameters of $\alpha = 0.1$ and $\gamma = \delta = 0.5$ for trend and seasonality, respectively. The model resulted in a Mean Absolute Percentage Error (MAPE) of 6.14, indicating a good prediction model result [13]. Furthermore, K.M.U.B. Konarasinghe [14] utilized Quadratic Trend Model, Double Exponential Smoothing (DES) techniques, and Holt's Winters three-parameter additive and multiplicative models to predict the outbreak in Indonesia and the Philippines. The study resulted in the DES technique being selected as the best model compared to the other two models. In another research, HW with additive or multiplicative model was applied for COVID-19 prediction in Sri Lanka. Both additive and multiplicative models gave good MAPE values which are 0.0207 and 0.2847, respectively. Table 1 summarizes the COVID-19 prediction models used by another researchers.

This paper is structured as follows: data collection, implemented prediction software, the study's applied model, and the metrics used to evaluate its effectiveness are summarised in Section 2. In Section 3, we present and discuss the study's findings. The work is summarised at the end of Section 4.

Table 1 - Prediction models on COVID-19

Authors, Year	Title	Method	Results
A. Sözen, A. D. Tuncer, and F. Kazancoglu, 2020	Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN, and LSTM approaches	ARIMA NARNN LSTM	LSTM was found to be the most accurate model (MAPE = 0.1640) in Switzerland
R. Sujath, Jyotir Moy Chatterjee, Aboul Ella Hassanien, 2020	A machine learning forecasting model for COVID-19 pandemic in India	Linear Regression, Multilayer perceptron, and Vector autoregression	The MLP technique outperforms the LR and VAR techniques in terms of prediction accuracy.

Ardabili et al., 2020	COVID-19 Outbreak Prediction with Machine Learning	MLP ANFIS	The models performed well without the assumptions typically used in epidemiological models.
Suzuki et al., 2020	Machine learning model estimating the number of COVID-19 infection cases overcoming 24 days in every province of South Korea (XGBoost and MultiOutputRegressor)	XGBoost MultiOutput Regressor	Accuracy = 82.4%
Mrutyunjaya Panda, 2020	Application of ARIMA and Holt-Winters forecasting model to predict the spreading of COVID-19 for India and its states	Holt's Winter ARIMA	Accuracy of Holt's Winter = 87.9% Accuracy of ARIMA (4,1,1) = 99.8%
Gothai et al., 2021	Prediction of COVID-19 growth and trend using machine learning approach	LR SVM Holt's Winter	Accuracy of Holt's Winter = 87.7%
I Djakaria and S E Saleh,2021	Covid-19 forecast using Holt-Winters exponential smoothing	multiplicative Holt-Winters exponential smoothing	Using trend and seasonality, smoothing parameters α = 0.1 and $\gamma = \delta = 0.5$, respectively, yield the best forecasting model with a MAPE value of 6.14.
K.M.U.B. Konarasinghe, 2021	Forecasting COVID -19 Outbreak in the Philippines and Indonesia	Quadratic Trend Model Double Exponential Smoothing (DES) techniques Holt's Winters three-parameter additive and multiplicative models	Daily infectious case forecasting in Indonesia and the Philippines is most accurately performed using DES.
S.S. Wickramasinghe and K.M.U.B. Konarasinghe, 2022	Forecasting COVID-19 Daily Infected Cases in Sri Lanka by Holt-Winters Model	Holt-Winters three parameters with additive or multiplicative models Holt-Winters three parameter with additive or multiplicative models	Multiplicative model:MAPE = 0.2847, MAD = 0.0187 and MSE=0.0005 Additive model:MAPE = 0.0207,MAD = 0.0187 and MSE=0.0005

2. Methods

There are four main components of methodology for this study which are the dataset used, software used, forecasting methods, and performance evaluation metrics used. Each part will be further discussed in the subsections.

2.1 Data Collection

COVID-19 data were obtained from a website Malaysian Ministry of Health. The number of confirmed COVID-19 cases reported in Malaysia on a daily basis was documented and saved in Microsoft Excel. The database contains 555 points between January 24th, 2020, and July 31st, 2021.

2.2 Implemented Software

In this study, the "Waikato Environment for Knowledge Analysis" (WEKA) software [15] was used to forecast daily instances for the following 14 days, from August 01st to August 14th, 2021. Microsoft Excel was used to organize and analyze the data.

2.3 Models for Prediction

In this paper, we compared two models: the LR model and HW model. LR is a fundamental Machine Learning approach often employed [16]. As a statistical approach, it may also be used for predictive research. In LR, a dependent variable is connected to one or more independent variables in a straight line, as the name suggests. Linear regression, which displays a straight line between the independent and dependent variables, estimates the dependent variable's value as it changes in response to the independent factors.[17]. In regression, as shown by Equation (1), y is correlated with x.

$$y = \beta_0 + \beta_1 x \qquad (1)$$

where β_0 is the y-intercept and β_1 is the pitch [18]. Optimizing the values of β_0 (intercept) and β_1 (coefficient) in a Linear Regression model yields a best-fit regression line.

Using HW forecasting, one may model and forecast the evolution of a set of variables through time. It is common to practice using HW, a well-known technique for predicting time series. Modeling time series data is what HW model does best. A time series can be expected using HW model, a method for modeling its average value, slope over time, and cyclical recurrence. There is a 0-1 range of values for these components. HW model is classified as multiplicative and additive based on the seasonal trend. When the seasonality impact is not related to the present middle point of the time series, an additive model such as HW is used. For seasonal effects that depend on the overall average of a time series, such as an increase in seasonal variation as the standard rises, HW propose a multiplicative model [19]. HW encrypt additional historical data to forecast potential conventional data using the exponential smoothing approach [17]. Triple exponential smoothing is another name for HW approach, which is simply a combination of three different smoothing techniques—Simple Exponential Smoothing (SES), Holt's Exponential Smoothing (HES), and Winter's Exponential Smoothing (WES)[20]. The model can be written in mathematical form as follows:

$$\hat{y}_{t+h} = a_t + hb_t + s_{t+h-p} \qquad (2)$$

where a_t, b_t, and s_t are given by

$$a_t = \alpha(y_t - s_{t-p}) + (1-\alpha)(a_{t-1} + b_{t-1}) \qquad (3)$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \qquad (4)$$

$$s_t = \gamma(y_t - a_t) + (1-\gamma)s_{t-p} \qquad (5)$$

where the series' level, slope, and seasonality at time *t* are denoted by a_t, b_t, and s_t, respectively. Additionally, *p* stands for the year's seasonal distribution. The smoothing parameters are represented by α, β, and γ, and the values range from 0 to 1. Meanwhile, *h* represents the duration during which the prediction is made [21]. The value of parameters used in this study is tabulated in Table 2 below.

Table 2 - Parameters value of HW model

α	0.2
β	0.2
γ	0.2
Year's seasonal distribution (*p*)	12

2.4 Model Evaluation

Throughout the training and testing phases, we only used one performance metric, MAPE, to evaluate the effectiveness of the methods used in this study. According to [22], MAPE is an effective indicator for use in forecasting. The MAPE metric determines how precise a given forecasting method is. This precision is expressed as a percentage and can be computed in Equation (6).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| \quad (6)$$

where A_i is the actual number of confirmed instances, F_i is the expected number of confirmed cases, and n is the total number of observations.

Since the MAPE is already in terms of the actual data's percentage error, the forecasting model's accuracy can be calculated by the difference between the MAPE value and 100%. The formula for the accuracy is as fellow:

$$Accuracy = 100\% - MAPE \quad (7)$$

3. Results and Discussion

Table 3 illustrates the summary output for regression analysis of both models. Regression analysis was performed to observe the relationship between the date (independent variable) and confirmed cases (dependent variable). In this table, we are interested in two components: the Multiple R and R^2. Multiple R represents the correlation coefficient and the values ranging from 0 (no correlation) to 1 (perfect correlation) [23]. For this study, the Multiple R values for the LR model and HW model are o.6863 and 0.7107, respectively. Both models give almost similar values of Multiple R, indicating a positive and robust correlation between date and confirmed cases [24].

Meanwhile, R^2 represents the coefficient of determination. R^2 indicates how well the regression model describes the observed data [25]. It also shows the proportion of variance in the dependent variable that the independent variable can explain. In this case, 0.4709 or 47.09% of the data fall within the LR line, and HW model can explain 0.5051 or 50.51% of the data. Fig. 1 visualizes the forecast value of the LR model for the next 14 days of daily confirmed COVID-19 cases in Malaysia. The orange line, which represents the forecasted value for the LR model, shows an increasing linear trend during the forecasting period. As for HW model, the last 14 points of the time showed a fluctuating trend, like the actual cases. The graph is shown in Fig. 2.

Table 3 - Regression analysis

	LR	**HW**
Multiple R	0.6863	0.7107
R^2	0.4709	0.5051
Adjusted R^2	0.4700	0.5042
Standard Error	3261.7744	2908.7907
Observations	569	569

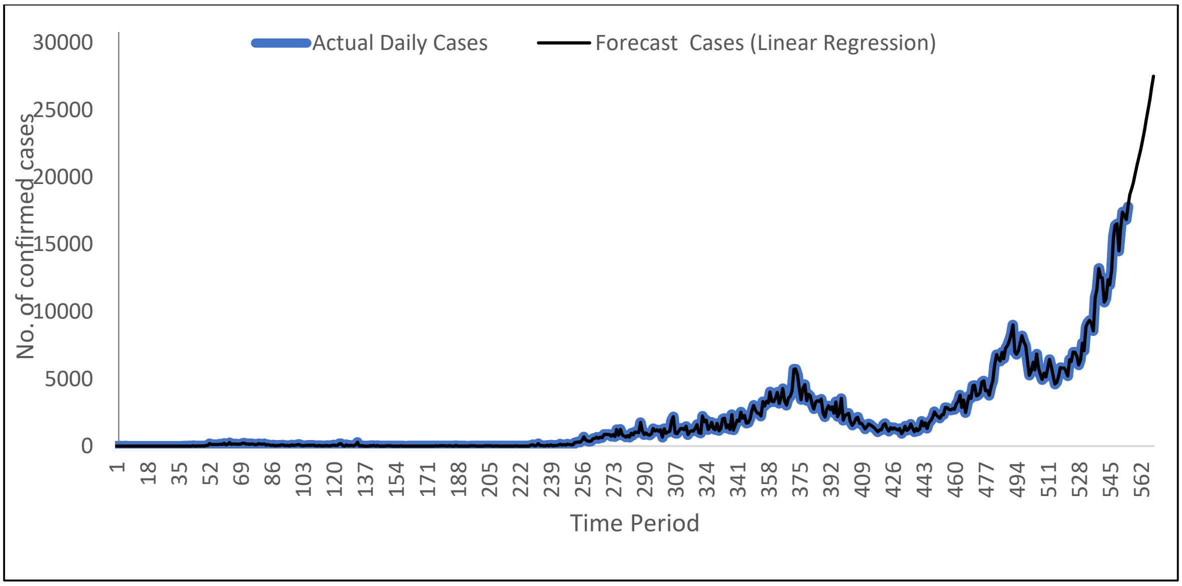

Fig. 1 - Forecasted cases of LR for 14-days ahead

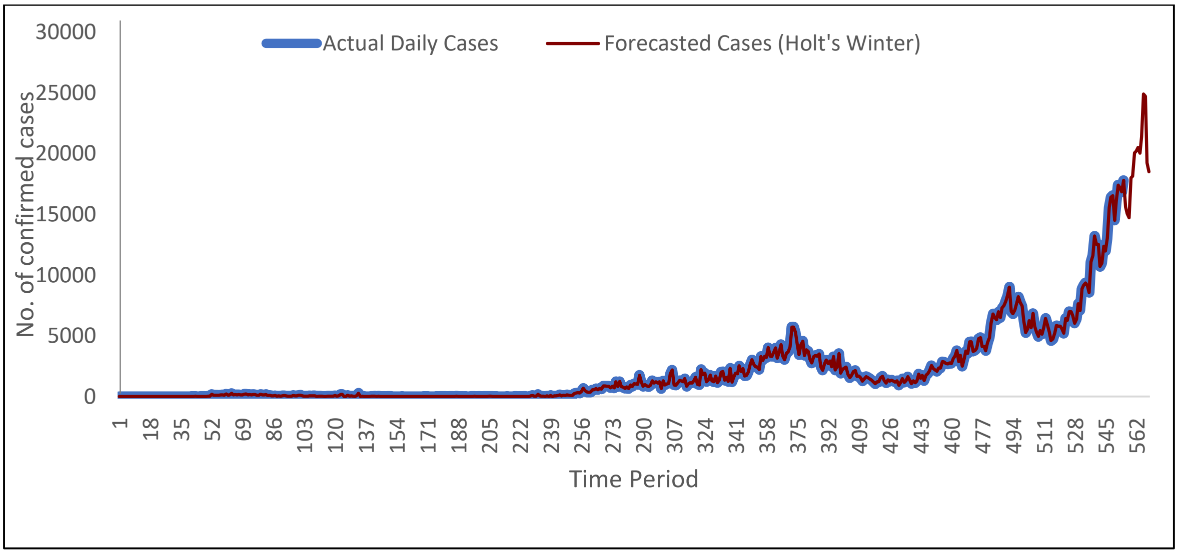

Fig. 2 - Forecasted cases of HW for 14-days ahead

The comparison of the cases for the LR model and HW model with the actual daily cases was visualized in Fig. 3.

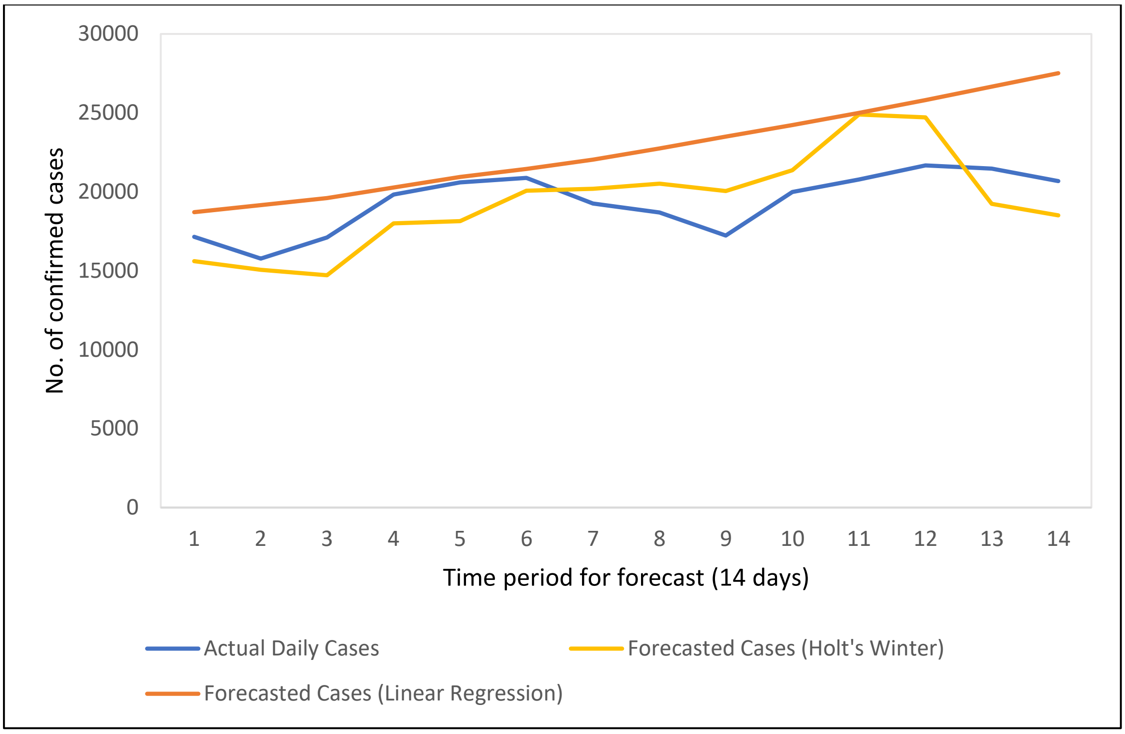

Fig. 3 - Comparison between actual and forecasted daily confirmed cases for LR and HW model

The findings of the comparison between the two models employed here are shown in Table 4. Both models performed admirably, with the LR model achieving an average accuracy of 82% and HW model achieving an average of 89%. In contrast to LR, HW model performed better in predicting daily COVID-19 cases in Malaysia. The MAPE metric also suggests that HW model is inferior to LR.

Table 4 - MAPE and Accuracy of LR and HW Model

Algorithms	MAPE	Accuracy
LR	17.30%	82.70%
HW	10.36%	89.64%

From the result, we could observe that the LR model can give a relatively high in terms of model prediction accuracy but still being outperformed by HW model. This was probably because of a few drawbacks of the LR model in predicting time series data. According to [25], LR's drawbacks include the fact that it frequently examines a correlation between the average of the input and output variables. Just as the mean does not adequately describe a single variable, linear regression does not provide a complete picture of variable connections. Furthermore, LR is a type of analysis that only looks at linear relationships between the variables that are being analyzed [26]. In other words, it assumes that there is a straight connection between them. This may not be optimum in this study, such that the COVID-19 data in Malaysia does not give a clear linear trend, and a single independent variable may not be sufficient to explain the dependent variable in many situations [27].

On the other hand, HW model can give higher accuracy since the mathematical model includes the trend and seasonality components, which correspond with the real-world COVID-19 data that has a lot of seasonality and trends [28]. In this study, the parameters for HW model were chosen by default in WEKA software. Therefore, an optimum value of the parameters can be done for future work in order to achieve higher accuracy for the model's prediction. Furthermore, in order to administer the health system adequately, it is essential to have an accurate short-term prognosis of the number of cases and deaths of infectious diseases like COVID-19. Because of this, this paper focused on predicting COVID-19 cases in the next 14 days. Allocating hospital beds and ventilators, deciding whether to put up more field hospitals, educating more health professionals, etc., may be better planned for if one knows how much strain the health system will be under the next day [29].

4. Conclusion

The purpose of this research is to determine which model can best forecast the future course of COVID-19 using available data. The LR and HW forecasting models were used using COVID-19 data from Malaysia to achieve this. Based on the results of this research, it can be concluded that HW model outperforms the LR model in terms of adequately predicting future confirmed instances. Analysis of historical data and projections of future trends show that the number of confirmed cases has stayed relatively constant. The created prediction model can be a handy tool for decision-making among public health. It can ensure resources are used effectively to reduce death and morbidity from a COVID-19 pandemic. It is possible to arrange preparation and control procedures to be implemented ahead of time. For future work, the optimum parameter can be determined for the prediction models since this study used the default setting of the parameters. Thus, the prediction accuracy of the model can be improved.

Acknowledgment

The authors would like to thank all collaborating partners for their support and Universiti Teknologi Malaysia for the sponsorship of the study under research grant no.: Q.K130000.3856.18J91.

References

[1] World Health Organization, "Novel Coronavirus (2019-nCoV) Situation Report-1," *WHO Bulletin*, 2020. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20a99c10_4.

[2] World Health Organization, "Novel Coronavirus (2019-nCoV) Situation Report-51," *WHO Bulletin*, 2020. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10.

[3] N. H. Abdullah, "Pengesanan Kes Baharu yang Disahkan Dijangkiti 2019 Novel Coronavirus (2019-nCoV) di Malaysia," *Kenyataan Akhbar KP KKM 25 Januari 2020*, pp. 1-3, 2020, [Online]. Available: http://covid-19.moh.gov.my/terkini/012020/situasi-terkini-25-jan-2020/04 KPK 25012020.pdf.

[4] Crisis Preparedness Response Centre Malaysia, "Situasi Semasa Pandemik COVID-19 Di Malaysia," *COVID-19 Malaysia 2020*, 2020. http://covid-19.moh.gov.my/ (accessed May 12, 2021).

[5] N. Talkhi, N. Akhavan Fatemi, Z. Ataei, and M. Jabbari Nooghabi, "Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: A comparison of time series forecasting methods," *Biomed. Signal Process. Control*, vol. 66, 2021, doi: 10.1016/j.bspc.2021.102494.

[6] İ. Kırbaş, A. Sözen, A. D. Tuncer, and F. Ş. Kazancıoğlu, "Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches," *Chaos, Solitons and Fractals*, vol. 138, 2020, doi: 10.1016/j.chaos.2020.110015.

[7] R. Sujath, J. Moy Chatterjee, and A. E. Hassanien, "A machine learning forecasting model for COVID-19 pandemic in India," *Stoch. Environ. Res. Risk Assess.*, vol. 34, no. 7, pp. 959-972, 2020, doi: 10.1007/s00477-020-01827-8.

[8] N. Alballa and I. Al-Turaiki, "Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review," *Informatics in Medicine Unlocked*, vol. 24. Elsevier, p. 100564, Jan. 01, 2021, doi: 10.1016/j.imu.2021.100564.

[9] S. F. Ardabili *et al.*, "COVID-19 Outbreak Prediction with Machine Learning," *Algorithms*, vol. 13, no. 10, p. 249, 2020, doi: 10.3390/a13100249.

[10] Y. Suzuki, A. Suzuki, S. Nakamura, T. Ishikawa, and A. Kinoshita, "Machine learning model estimating number of COVID-19 infection cases over coming 24 days in every province of South Korea (XGBoost and MultiOutputRegressor)," *medRxiv*, no. Ml, pp. 1-11, 2020, doi: 10.1101/2020.05.10.20097527.

[11] E. Gothai, R. Thamilselvan, R. R. Rajalaxmi, R. M. Sadana, A. Ragavi, and R. Sakthivel, "Prediction of COVID-19 growth and trend using machine learning approach," *Mater. Today Proc.*, Apr. 2021, doi: 10.1016/J.MATPR.2021.04.051.

[12] M. Panda, "Application ofARIMA and Holt-Winters forecasting model to predict the spreading ofCOVID-19 for India and its states," 2020.

[13] D. I and S. S E, "Covid-19 forecast using Holt-Winters exponential smoothing," *J. Phys. Conf. Ser.*, 2021, doi: 10.1088/1742-6596/1882/1/012033.

[14] K. K.M.U.B, "Forecasting COVID -19 Outbreak in the Philippines and Indonesia," *J. New Front. Healthc. Biol. Sci.*, vol. 2, no. 1, pp. 1-19, 2021.

[15] E. Frank, M. A. Hall, and I. H. Witten, "The WEKA workbench," *Data Min.*, pp. 553-571, 2017, doi: 10.1016/b978-0-12-804291-5.00024-6.

[16] P. Sarkar, "What is Linear Regression in Machine Learning," *knowledgehut*, 2019. https://www.knowledgehut.com/blog/data-science/linear-regression-for-machine-learning (accessed Aug. 19, 2022).

[17] E. Gothai, R. Thamilselvan, R. R. Rajalaxmi, R. M. Sadana, A. Ragavi, and R. Sakthivel, "Prediction of COVID-

19 growth and trend using machine learning approach," *Mater. Today Proc.*, Apr. 2021, doi: 10.1016/j.matpr.2021.04.051.
[18] R. K. Mojjada, A. Yadav, A. V. Prabhu, and Y. Natarajan, "Machine learning models for covid-19 future forecasting," *Mater. Today Proc.*, Dec. 2020, doi: 10.1016/j.matpr.2020.10.962.
[19] H. Swapnarekha, H. S. Behera, J. Nayak, B. Naik, and P. S. Kumar, "Multiplicative Holts Winter Model for Trend Analysis and Forecasting of COVID-19 Spread in India," *SN Comput. Sci.*, vol. 2, no. 5, Sep. 2021, doi: 10.1007/S42979-021-00808-0.
[20] L. Ariton, "A Thorough Introduction to Holt-Winters Forecasting," *Analytics Vidhya*, 2021. https://medium.com/analytics-vidhya/a-thorough-introduction-to-holt-winters-forecasting-c21810b8c0e6 (accessed Sep. 13, 2021).
[21] A. M. Awajan, M. T. Ismail, and S. AL Wadi, "Improving forecasting accuracy for stock market data using emd-hw bagging," *PLoS One*, vol. 13, no. 7, p. e0199582, Jul. 2018, doi: 10.1371/journal.pone.0199582.
[22] U. Khair, H. Fahmi, S. Al Hakim, and R. Rahim, "Forecasting Error Calculation with Mean Absolute Deviation and Mean Absolute Percentage Error," in *Journal of Physics: Conference Series*, 2017, vol. 930, no. 1, doi: 10.1088/1742-6596/930/1/012002.
[23] K. Kumari and S. Yadav, "Linear regression analysis study," *J. Pract. Cardiovasc. Sci.*, vol. 4, no. 1, p. 33, 2018, doi: 10.4103/JPCS.JPCS_8_18.
[24] D. Mindrila and M. E. Phoebe, "Scatterplots and Correlation."
[25] A. Schneider, G. Hommel, and M. Blettner, "Linear Regression Analysis: Part 14 of a Series on Evaluation of Scientific Publications," *Dtsch. Arztebl. Int.*, vol. 107, no. 44, p. 776, Nov. 2010, doi: 10.3238/ARZTEBL.2010.0776.
[26] P. Floam, "The Disadvantages of Linear Regression," 2013. https://sciencing.com/disadvantages-linear-regression-8562780.html (accessed Apr. 30, 2022).
[27] S. Bzovsky *et al.*, "The clinician's guide to interpreting a regression analysis," 2022, doi: https://doi.org/10.1038/s41433-022-01949-z.
[28] B. Snehal, "Holt Winter's Method for Time Series Analysis | Holt Winter's Method," 2021. https://www.analyticsvidhya.com/blog/2021/08/holt-winters-method-for-time-series-analysis/ (accessed Apr. 30, 2022).
[29] M. C. Medeiros, A. Street, D. Valladao, G. Vasconcelos, and E. Zilberman, "Short-Term Covid-19 Forecast for Latecomers," 2021.