Indonesian Journal of Electrical Engineering and Computer Science
Vol. 31, No. 1, July 2023, pp. 299~304
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v31.i1.pp299-304

Comparison of machine learning algorithms with regression analysis to predict the COVID-19 outbreak in Thailand

Supanee Sengsri[1], Kheamparit Khunratchasana[2]
[1]Department of Education Technology and Communication, Faculty of Education, Naresuan University, Phitsanulok, Thailand
[2]Department of Computer Science, Faculty of Science and Technology, Phetchabun Rajabhat University, Phetchabun, Thailand

Article Info

Article history:

Received Aug 29, 2022
Revised Feb 3, 2023
Accepted Feb 12, 2023

Keywords:

COVID-19
Decision tree
K-nearest neighbor
LASSO
Machine learning
Random forest
XGBoost

ABSTRACT

Coronavirus disease (COVID-19) is a public health problem in Thailand. Currently, there are more than 5 million infected people and the rate has been increasing at some point. It is therefore important to forecast the number of new cases over a short period of time to assist in strategic planning for the response to COVID-19. The purpose of this research paper was to compare the efficiency and prediction of the number of COVID-19 cases in Thailand using machine learning of 8 models using a regression analysis method. Using the 475-day dataset of COVID-19 cases in Thailand, the results showed that the predictive accuracy model (R2 score) from the testing dataset was the random forest (RF) model, which was 99.06%, followed by K-nearest neighbor (KNN), XGBoost. And the decision tree (DT) had the precision of 98.97, 98.67, and 98.64, respectively. And the results of the comparison of the number of infected people obtained from the prediction The models that predicted the number of real infections were the decision tree, random forest, and XGBoost, which were effective at predicting the number of infections correctly in the 2-4 day period.

Corresponding Author:

Kheamparit Khunratchasana
Department of Computer Science, Faculty of Science and Technology, Phetchabun Rajabhat University
Phetchabun 67000, Thailand
Email: kparit1.khu@pcru.ac.th

1. INTRODUCTION

Thailand is a country in Southeast Asia and is part of association of Southeast Asian nations (ASEAN). Currently, it has been affected by the coronavirus disease (COVID-19) epidemic in various fields such as economy, public health, and people's daily living. Thailand finds the first individual infected with COVID-19, on 12 January 2020, the individual is a Chinese female tourist [1]. Currently, the total number of infections since the outbreak began in early 2020 has exceeded 4 million; the cumulative death toll has exceeded 28,500, but with the death toll, many individuals have been cured (as of April 30, 2022). COVID-19 had spread from Hubei, Wuhan City, China, around December 2019, and later spread to other cities across China and around the world.

The international committee on virus classification has given the official name COVID-19, and is derived from coronavirus disease 2019. The World Health Organization has declared the COVID-19 outbreak a Public Health Emergency of International Concern [2]–[4]. Coronavirus is a multi-strain virus caused by birds and mammals. It is a respiratory virus that can be fatal. But a healthy patient will recover without treatment [5], [6]. COVID-19 affects global citizens it presents various challenges to humanity. Researchers of various fields are trying to contribute to the fight against this epidemic through new ways by applying technologies such as artificial intelligence, and cloud computing [7].

From the information about the spread of COVID-19 around the world right now, artificial intelligence can be applied to create models to predict the spread of pathogens. By developing an artificial intelligence system, the goal is to develop the system to have intelligent behavior similar to that of humans [8]. Machine learning is a sub domain of Artificial Intelligence divided into three main categories: supervised learning, unsupervised learning, and reinforced learning, where algorithms can be used to predict the spread of COVID-19 [9], [10]. It can be applied to solve complex problems that arise in the real world. It has been applied by the application of machine learning in various fields, such as public health, autonomous vehicles, games, and robotics [11], [12]. Therefore, this research article has an idea that can be applied in public health in Thailand. The purpose of this research was to compare the efficiency and prediction of the number of COVID-19 cases in Thailand using machine learning by regression analysis and using 8 predictive models.

2. METHOD

For this research paper, 7 steps are used as:

Step 1: Data gathering, for the data set of the number of people infected with COVID-19 used in machine learning algorithm modeling to predict the number of infections. Using information from the Department of Disease Control, Thailand, this is published through the government's open information center [13]. Its main storage attributes are no, age, sex, nationality, notification date, and announce date.

Step 2: Data pre-processing, data preparation for model training and model testing was based on 476 days of daily reported number of infections reported from January 1, 2021 to April 20, 2022, and data from April 21-30, 2022 were used for comparison. The result of the prediction the total number of infected people is 4,077,415, with an average daily infection rate of 8584.32, representing 53.52% males and 46.48% females. The daily infected data is stored in csv files, consisting of 2 columns: date and announce date. Using the data to train the model and test the performance of the machine learning model.

Step 3: Choosing a model, for this research paper, a machine learning model belonging to the supervised learning category was chosen as a popular algorithm. It can be used to analyze regression or classification of data by correlating old data with new data [14]. Each machine learning model has the following details:

a. Linear regression (LR) is an algorithm that uses regression modeling to determine the relationship between independent and dependent variables and for forecasting. Linear regression is a statistical technique to create the most applicable regression models for predictive analysis in machine learning. The linear regression is shown in (1) [15].

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{1}$$

b. Polynomial regression (PR) is an algorithm that is suitable for independent and dependent variables with non-linear relationships [16]. The independent variable must be assigned to the nth degree polynomial of the dependent variable. The polynomial regression equation is shown in (2) [17].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 {x_1}^2 + \beta_3 {x_1}^3 + \ldots + \beta_n {x_1}^n \tag{2}$$

c. K-nearest neighbor (K-NN) regression is an algorithm that finds regression or classification by finding the relationship between old data and new dataA value must be given to parameter K, provided that K is not greater than the data to be calculated and the value of K must be odd and greater than 1. For the nearest neighbor method, this paper uses the Euclidean function. Shown as (3) [18].

$$D(p,q) = \sqrt{\textstyle\sum_{i=1}^{n}(q_i - p_i)^2} \tag{3}$$

Where D is the distance between p and q, n is the total data; i is the sequence of data.
And forecasting the upcoming data can be done by averaging the results from the position nearest to the point to be searched by the number of values of K shown in (4).

$$y = \frac{1}{K}\textstyle\sum_{i=1}^{k} y_i \tag{4}$$

Where y is the predicted result, K is the constant; y_i is the value of y in the i position.

d. Support vector regression (SVR) is a popular algorithm for classifying problems, but it can also be used in regression analysis. The SVR has a kernel that acts as the engine for analyzing the data. The principle of operation is to take the input data as an input vector and an output variable. After that select the kernel that is suitable for the data to be analyzed. The kernel defines a line separating the data clusters, called a hyperplane, that divides the data into two clusters equidistant from each edge of the cluster, with two borders each passing through each cluster's data point. Lines parallel to the hyperplane are called margin boundaries. They have two sides, positive and negative. Points on the margin boundary line are called support vectors [19]–[21].

e. Least absolute selection and shrinkage operator (LASSO) is a model used for regression analysis for high-dimensional data. It is used to optimize data and select the best features from over-minimization. It is a method of estimating parameters to be input into estimation models in regression analysis using the penalty function by squeezing most coefficients to zero. It is a linear regression technique which uses shrinkage. The shrinkage process makes LASSO better and more accurate and reduces errors by LASSO regression [22]. The equation for reducing the parameters is shown in (5) [15].

$$\sum_i^n \left(y_i - \sum_j x_{ij}\,\beta_j\right)^2 + \lambda \sum_{j=1}^{p} \left|\beta_j\right| \tag{5}$$

f. Decision tree is an algorithm based on decision-making principles with an inverted binary tree. It builds a model to predict the value of a target variable by learning simple decision rules inferred from data attributes. The survey will go through each branch with conditional division. The prediction is based on the leaf node, the top of which is the root, and the bottom that can't be branched is the leaf. It is determined from the starting point called the root note. If the data found meets the decision condition, it runs to the left of the root note to the point called the child node. A child node is considered to be terminated by an end point called a leaf node, wherein the data set used for training is divided into hierarchies [23].

g. Random forest (RF) is an algorithm that has improved capabilities over the decision tree model. Its working principle is to combine small-divided decision trees through a re-sampling process known as bagging. Multiple decision trees are generated by bootstrap re-sampling with substitution. Each node of the tree is extracted using a randomly selected subset of attributes for each tree. The results were divided into two types. If it is grouped, the result will be predicted by means of voting. If it is a regression analysis, the result will be predicted by means of finding the mean [24]–[26].

h. The XGBoost model is a machine learning algorithm that trains multiple decision trees to make the model more efficient. Accurate predictions can be made and at the same time the model shows a ranking of the input features where each decision tree learns from the tolerance of the previous one. As a result, the accuracy of predictions increases over time, and the model stops learning when the error values from the previous decision tree run out. This model also offers other benefits such as reduced run time by parallel and distributed computation. effectively dealing with missing values according to Mehta *et al*. [27] and Fang *et al*. [28].

Step 4: Model training, this step will take the data prepared from step 2.2 into 2 sets, namely training data set and testing data set, with a ratio of 80:20 by training the model to learn. Learn from python programming tutorial data and run a set of instructions from the scikit-learn library to determine predictive performance and apply 8 predictive models to predict the spread of COVID-19. It is cloud-processed with Google colaboratory via Jupyter notebooks, an efficient and free system, and uses pandas and numpy to manipulate time-series data, diagram data using matplotlib and sebon libraries [29].

Step 5: Evaluating the model, in this paper, the efficiency of the machine learning model was performed using the functions of the scikit-learn library by evaluating the accuracy of the predictions with R2scores and calculating the prediction error from the mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE). The equations for accuracy and error are shown in (6)-(9) [15].

$$R^2 = \frac{\text{Varianceexplainedbymodel}}{\text{Totalvariance}} \tag{6}$$

$$MAE = \frac{1}{n}\sum_j^n \left|y_j - \hat{y}_j\right| \tag{7}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{9}$$

Step 6: Hyperparameter tuning, different machine learning models have different unique parameters to control the model training for the best accuracy and optimum performance. Some models call the GridSearchCV function to find the appropriate parameters to test the model. In this research, the appropriate parameters for the model were as follows: polynomial regression set degree=7, K-NN set K=3, SVR set C=4000 and gamma=0.01, random forest set n_estimators=100, LASSO set alpha=0.01, max_iter=100, random_state=100, tol=0.001.

Step 7: Prediction, after the model has been created and the hyperparameter has been tuned to the model, it makes forecasting accuracy the highest value for each model. After that, the model was used to predict the number of people infected with COVID-19 for 10 days and compare the results with the actual number of infected people on April 21-30, 2022.

3. RESULTS AND DISCUSSION

The purpose of this research paper was to compare the efficiency and prediction of the number of COVID-19 cases in Thailand using machine learning by analyzing 8 regression models: linear regression, polynomial regression, k-nearest neighbor, support vector regression, LASSO, decision tree, random forest, and XGBoost. According to preliminary data processing, there were 475 days of COVID-19 cases in Thailand between January 1, 2021, and April 20, 2022. The number of infections gradually increased during the first 170 days, then increased exponentially until day 220, after which it gradually decreased and the number of infections increased again. Because the pathogen has a mutant variant named omicron. The variant can spread faster than the delta species, resulting in a leap in the number of infections. The leap in infections caused the government to take measures to regulate the people; and this action in turn caused the number of infected people to decrease. The trend of COVID-19 infections in Thailand is shown in Figure 1.

The results of the development of a machine learning model to compare the effectiveness of predicting the number of COVID-19 cases in Thailand by regression analysis. The model with the highest predictive efficiency was random forest with an R^2scores of 99.06% from the testing dataset, followed by K-NN, XGBoost and decision tree with R^2scores of 98.97, 98.67, and 98.64, respectively. The MAE and RMSE values were similar, consistent with the research of Bhadana [5], which used seven predictive models. The most efficient model was the decision tree, followed by random forest and polynomial regression. They had R^2scores of 100.00, 99.90, and 98.65, respectively. The prediction efficiency and error of each model are shown in Table 1.

The comparison of the number of infected people predicted by different models compared to the actual number of infected people, which on April 21-23, 2022, was about 20,000 people per day, after which the number of infected people decreased. As for the predicted value, it was found that the number of infected people was close to the actual number of people infected during April 21-23, 2022. The 3 models that were most similar to the actual number of infections were decision tree, random forest and XGBoost, had predicted values of 20,455, 19383.22, respectively. From the observations, it was found that the predicted values were from the 25th day onwards. Every model has an increased predictive effect as opposed to reality. The predicted values are shown in Table 2.

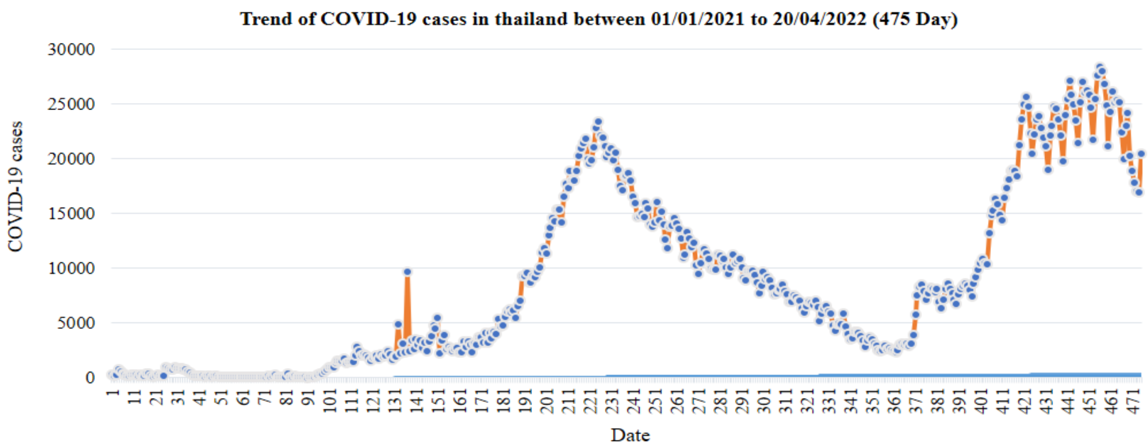

Figure 1. Trend graph of COVID-19 cases in Thailand between January 1, 2021 and April 20, 2022

Table 1. The predictive efficiency of each machine learning model

Algorithm	R^2 score		MAE	MSE	RMSE
	Training dataset	Testing dataset			
LR	50.43	49.79	4538.96	35140288.12	5927.92
PR	92.29	90.46	1840.81	6674894.20	2583.58
K-NN	99.35	98.97	576.18	723903.24	850.82
SVM	98.54	98.31	715.10	1183154.42	1087.73
LASSO	50.43	49.79	4538.96	35140288.24	5927.92
Decision Tree	100.00	98.64	645.15	949682.28	974.51
Random Forest	99.80	99.06	579.12	658359.49	811.39
XGBoost	99.37	98.67	685.03	932815.53	965.82

Table 2. Compare the actual number of COVID-19 cases in Thailand with the predicted results with a machine learning model

Date	Actual case	LR	PR	LASSO	K-NN	SVM	Decision tree	Random forest	XGboost
21/04/2022	21931.00	17951.93	13516.31	17951.93	18373.67	14660.06	20455.00	19383.22	19874.66
22/04/2022	21808.00	17992.15	12297.31	17992.15	18373.67	13958.34	20455.00	19383.22	19874.66
23/04/2022	20052.00	18032.38	11007.95	18032.38	18373.67	13287.50	20455.00	19383.22	19874.66
24/04/2022	17784.00	18072.60	9646.11	18072.60	18373.67	12652.92	20455.00	19383.22	19874.66
25/04/2022	14994.00	16906.14	26568.42	16906.14	25502.00	25173.29	25821.00	25424.03	25347.39
26/04/2022	13816.00	16946.36	26555.67	16946.36	24044.67	25157.38	24635.00	24730.95	24915.82
27/04/2022	14887.00	16986.58	26515.03	16986.58	23900.67	25132.25	21678.00	22922.81	24214.16
28/04/2022	14437.00	17026.80	26445.28	17026.80	24875.67	25095.81	25389.00	24900.10	25311.95
29/04/2022	14053.00	17067.03	26345.16	17067.03	24875.67	25045.62	27560.00	26728.55	26088.02
30/04/2022	12888.00	17107.25	26213.38	17107.25	26596.33	24978.79	27560.00	27104.42	26088.02

4. CONCLUSION

The world has been affected by the COVID-19 outbreak, which has caused worldwide concern. In this research paper, we applied a machine learning model to predict the spread of COVID-19 in Thailand, which has a different pattern of transmission within the country than other countries. Outgoing data is processed in the format of the date and number of outbreaks for each day. The results of this study revealed that the most effective model for prediction was random forest, with 99.06% predictive efficiency from the testing data set, followed by K-NN, XGBoost and decision tree, which had the prediction accuracy of 98.97, 98.67, and 98.64, respectively. From the comparison of the actual number of infections with the predicted values, it was found that each model had a different predictive ability close to the actual value. Models that have predictive results close to the actual number of infected people are decision tree, random forest, and XGBoost, which are effective at predicting accurate pre-infection numbers in a short period of 2-4 days. After that, the predicted value will increase as opposed to the actual situation. It may be a result of the volume of the epidemic increasing and decreasing according to the situation. Including various measures that the government has announced a prevention policy and depends on the cooperation of citizens in Thailand. Based on the development of a machine learning model for predicting COVID-19 cases, it can be concluded that no specific model is the best because the data used to make predictions, because it depends on the nature of the epidemic in each country, including the cooperation of the citizens of that country and the COVID-19 virus is an organism that can reproduce on its own to survive. Therefore, multiple models must be combined for prediction in order to achieve the prediction results as close to the actual as possible.

ACKNOWLEDGEMENTS

Thank you to the Department of Computer Science, Faculty of Science and Technology, Research and Development Institute, and Phetchabun Rajabhat University for providing support in terms of time and resources related to this research.

REFERENCES

[1] A. Ratnarathon, "Coronavirus infectious disease-2019 (COVID-19): a case report, the first patient in Thailand and outside China," vol. 14, no. 2, pp. 116–123, 2020.
[2] N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "COVID-19 epidemic analysis using machine learning and deep learning algorithms," *MedRxiv*, 2020.
[3] S. A. Kumar, H. Kumar, V. Dutt, and P. Dixit, "Deep Analysis of COVID-19 Pandemic using Machine Learning Techniques," *Global Journal on Innovation, Opportunities and Challenges in Applied Artificial Intelligence and Machine Learning [ISSN: 2581-5156 (online)]*, vol. 4, no. 2, 2020.
[4] M. Chakraborty, M. S. Mahmud, T. J. Gates, and S. Sinha, "Analysis and prediction of human mobility in the United States during the early stages of the COVID-19 pandemic using regularized linear models," *Transportation Research Record*, p. 03611981211067794, 2022.
[5] V. Bhadana, A. S. Jalal, and P. Pathak, "A Comparative Study of Machine Learning Models for COVID-19 prediction in India," in *2020 IEEE 4th Conference on Information & Communication Technology (CICT)*, 2020, pp. 1–7.
[6] M. Pourhomayoun and M. Shakibi, "Predicting mortality risk in patients with COVID-19 using artificial intelligence to help medical decision-making," *MedRxiv*, 2020.
[7] M. G. Al-Hamiri, H. F. Abdulsada, and L. A. Abdul-Rahaim, "Applications of artificial intelligence with cloud computing in promoting social distancing to combat COVID-19," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 3, pp. 1550–1556, 2021.
[8] H. B. Syeda *et al.*, "Role of machine learning techniques to tackle the COVID-19 crisis: Systematic review," *JMIR medical informatics*, vol. 9, no. 1, p. e23811, 2021.
[9] S. Kekre, "Comparative Study and Proposed Approach for Multi-Variate Regression through Gradient Boosting," 2020.
[10] W. Shafik, S. M. Matinkhah, S. S. Afolabi, and M. N. Sanda, "A 3-dimensional fast machine learning algorithm for mobile unmanned aerial vehicle base stations," *Int J Adv Appl Sci ISSN*, vol. 2252, no. 8814, p. 8814, 2020.

[11] M. Maleki, M. R. Mahmoudi, D. Wraith, and K.-H. Pho, "Time series modelling to forecast the confirmed and recovered cases of COVID-19," *Travel medicine and infectious disease*, vol. 37, p. 101742, 2020.
[12] R. Sujath, J. M. Chatterjee, and A. E. Hassanien, "A machine learning forecasting model for COVID-19 pandemic in India," *Stochastic Environmental Research and Risk Assessment*, vol. 34, pp. 959–972, 2020.
[13] Disease Control, Thailand, "Daily COVID-19 Report Thailand Information." 2022. Accessed: Apr. 20, 2022. [Online]. Available: https://data.go.th/dataset/covid-19-daily
[14] M. R. Romadhon and F. Kurniawan, "A Comparison of Naive Bayes Methods, Logistic Regression and KNN for Predicting Healing of Covid-19 Patients in Indonesia," in *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, 2021, pp. 41–44.
[15] F. Rustam *et al.*, "COVID-19 future forecasting using supervised machine learning models," *IEEE access*, vol. 8, pp. 101489–101499, 2020.
[16] R. Sadik, M. L. Reza, A. Al Noman, S. Al Mamun, M. S. Kaiser, and M. A. Rahman, "COVID-19 Pandemic: A Comparative Prediction Using Machine Learning," *International Journal of Automation, Artificial Intelligence and Machine Learning*, vol. 1, no. 1, pp. 1–16, 2020.
[17] H. Singh and S. Bawa, "Predicting COVID-19 statistics using machine learning regression model: Li-MuLi-Poly," *Multimedia Systems*, vol. 28, no. 1, pp. 113–120, 2022.
[18] G. Pattnaik and K. Parvathy, "Machine learning-based approaches for tomato pest classification," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 2, pp. 321–328, 2022.
[19] L. M. Abouelmagd, A. Darwish, and A. E. Hassanien, "Impact of COVID-19 Pandemic on Diet Prediction and Patient Health Based on Support Vector Machine," *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2021*, vol. 1339, p. 64, 2021.
[20] H. Herlawati, "COVID-19 spread pattern using support vector regression," *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, vol. 8, no. 1, pp. 67–74, 2020.
[21] A. Salama, A. Darwsih, and A. E. Hassanien, "Artificial intelligence approach to predict the covid-19 patient's recovery," *Digital Transformation and Emerging Technologies for Fighting COVID-19 Pandemic: Innovative Approaches*, pp. 121–133, 2021.
[22] D.-L. Chen, J.-H. Cai, and C. C. Wang, "Identification of key prognostic genes of triple negative breast cancer by LASSO-based machine learning and bioinformatics analysis," *Genes*, vol. 13, no. 5, p. 902, 2022.
[23] P. Theerthagiri, I. J. Jacob, A. U. Ruby, and Y. Vamsidhar, "Prediction of COVID-19 Possibilities using KNN Classification Algorithm," 2020.
[24] 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, p. 100564, 2021.
[25] N. Ayala, A. Monleon-Getino, J. Canela-Soler, and T. Chadwick-Lobos, "Random forest using smartphone GPS in first wave of COVID-19 in the Maule region, Chile," 2023.
[26] A. R. Thatipalli, P. Aravamudu, K. Kartheek, and A. Dennisan, "Exploring and comparing various machine and deep learning technique algorithms to detect domain generation algorithms of malicious variants," *Computer Science and Information Technologies*, vol. 3, no. 2, pp. 94–103, 2022.
[27] M. Mehta, J. Julaiti, P. Griffin, and S. Kumara, "Early stage machine learning–based prediction of US county vulnerability to the COVID-19 pandemic: machine learning approach," *JMIR public health and surveillance*, vol. 6, no. 3, p. e19446, 2020.
[28] Z. Fang, S. Yang, C. Lv, S. An, and W. Wu, "Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: a time-series study," *BMJ open*, vol. 12, no. 7, p. e056685, 2022.
[29] H. H. Holm, A. R. Brodtkorb, and M. L. Sætra, "GPU computing with Python: Performance, energy efficiency and usability," *Computation*, vol. 8, no. 1, Art. no. 1, 2020.

BIOGRAPHIES OF AUTHORS

Supanee Sengsri is an Associate Professor in Educational Technology at Faculty of Education, Naresuan University, Thailand. She completed her M.Ed. and Ph.D. degrees in Educational Technology and Communication from Chulalongkorn University, Thailand. In 2006, she received the UMAP Scholarship to train in Distance Education from Madras University and e-learning design and development from Indian Institute of Information Technology, Allahabad India. Her research interests include hybrid flexible (HyFlex) learning and instruction, professional teacher training in educational technology, Content and Language integrated Pedagogy via electronic communication (e-CLIP), and highly interactive web-based learning and application artificial intelligence and machine learning in education. She can be contacted at email: supanees@nu.ac.th.

Kheamparit Khunratchasana is an Assistant Professor of Information Technology. He has been working as lecturer in Computer Science at the Faculty of Science and Technology, Phetchabun Rajabhat University, Thailand. He wrote textbooks on Python programming to teach in higher education. He holds a Master's degree in Technology and Communication; a minor in Computer Studies, from Naresuan University, Thailand, and graduated with a bachelor's degree at King Mongkut's Institute of Technology North Bangkok, Bangkok, Thailand. He has interests in artificial intelligence, machine learning, deep learning, and data science. He can be contacted at email: kparit1.khu@pcru.ac.th.