Journal of Nonlinear and Convex Analysis
Volume 26, Number 7, 2025, 2235–2244

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

NEURAL NETWORK MULTI-ALGORITHM COMBINATION MODEL FOR EPIDEMIC PREDICTION

SHAOJUAN MA, XINYU BAI*, ZHENG HAO, AND HONGLIN CHEN

Abstract. The statistical models to effectively predict coronavirus disease 2019 (COVID-19) hold immense practical significance globally. This study aimed to introduce a multi-algorithm combination model to explore COVID-19. First, this study constructed the hybrid model for Convolutional neural network (CNN)-Bidi rectional long short term memory (BiLSTM)-Attention. Second, the COVID-19 data from Europe were selected for experimentation after denoising using empirical mode decomposition. Finally. the prediction accuracy of the proposed method was compared with different methods, to verify its high accuracy from a statistical standpoint.

1. INTRODUCTION

The novel coronavirus disease 2019 (COVID-19) has spread rapidly globally since its outbreak, posing a significant threat to global health even a few months after its outbreak [11, 19]. The World Health Organization has declared it a global health emergency due to its rapid spread and high mortality rate [1, 9]. The recurrence of the epidemic has had a serious impact on normal development and people's safety [2, 4]. Hence, forecasting new cases is crucial for allocating medical resources and protecting people's lives.

Two methods are used for predicting the new cases of COVID-19. One involves making predictions by discussing infectious disease models. Zhou et al. estimated the epidemic regeneration number and predicted the spread of the infectious disease using the susceptible-exposed-infected-removed (SEIR) model [26]. Wu et al. simulated and predicted COVID-19 in China and globally. They concluded that, if left unchecked, the number of COVID-19 infections might increase exponentially in several major cities in China [23]. Chan et al. suggested vigilant prevention and control behavior because of the uncertainty of incubation period and asymptomatic infections [10]. The susceptible-exposed-asymptomatic-infected-removed-dead (SEAIRD) model was proposed by Liu et al. to integrate asymptomatic infected and dead patients [12].

Another method involves prediction using statistical methods. Usually, this method estimates COVID-19 mainly using the autoregressive integrated moving average (ARIMA) model [3] and the seasonal ARIMA model [5]. Yang et al. predicted Italian data using the ARIMA model [24]. The multilayer perceptron neural

---

2020 *Mathematics Subject Classification.* 62M20.

Key words and phrases. BiLSTM, COVID-19, empirical mode decomposition, hybrid model.

This work was supported by the grants from the National Natural Science Foundation of China (No.12362005), Ningxia higher education first-class discipline construction funding project (NXYLXK2017B09), Major Special project of North Minzu University (No.ZDZX201902).

*Corresponding author.

network (MLP)-ARIMA model was proposed to estimate the daily number of new patients with COVID-19 globally [13]. With the advancement of machine learning, many scholars have applied the aforementioned model to predict COVID-19. Siti et al. applied the support vector regression (SVR) model to estimate the number of deaths and cases in Malaysia [21]. Parbat et al. constructed an SVR model with a radial basis function kernel to predict daily, recovery, and death cases [17].

Deep learning, a subset of machine learning, performs better in nonlinear prediction. Long short-term memory (LSTM) is an algorithm developed from deep learning. In 1997, Schmidhuber and Hochreiter improved the cyclic neural network and proposed the LSTM model, promoting the development of the cyclic neural network family and expanding its application range [7]. Nguyen et al. used the LSTM network to calculate the cases for the respiratory diseases in central Taiwan [15]. Shashank constructed K-mean-LSTM to predict short-term COVID-19 cases in Louisiana [20]. Luo established a prediction model of daily confirmed infections with US time-series data based on LSTM and XGBoost algorithms [14]. Farah used LSTM and BiLSTM to forecast the cases, recovery, and deaths of COVID-19 in 10 countries [6]. Unlike LSTM, BiLSTM could better explore the temporal features within the data. CNN decreased the number of weights to be trained and retained most of the important information, reducing the complexity of the model and effectively improving the efficiency of neural network model training.

Traditional denoising methods are widely used to solve the strong noise problem of complex time series [25], such as Fourier transform (FT) [18], singular spectrum analysis [22], and wavelet analysis [16]. Compared with traditional methods, empirical mode decomposition (EMD) has obvious advantages in processing complex time-series data [8]. It can decompose complex time series into signal components at different frequencies.

Given the problems of data noise disturbance and deep neural network application, this study applied EMD to the model to propose the EMD-CNN-BiLSTM-Attention model to estimate the daily newly confirmed COVID-19 cases in Europe. In this study, a model was built for optimizing Variational mode decomposition (VMD) with singular spectrum analysis (SSA) by establishing a new comprehensive fitness function. Then, a hybrid noise reduction model was established through Pearson correlation coefficient (PCC). Finally, we verified the theoretical model with real time-series data.

2. Ombination of deep neural network models

2.1. **Single deep learning model.**

Convolutional neural network (CNN) is an effective feedforward neural network for feature extraction and pattern recognition. It can be originally traced back to the 1960s. The neurons within a CNN can respond to the surrounding neurons. The essence of CNN is a multi-layer perceptron with weight-sharing connection mode and local connection, which can significantly decrease the number of weights to be trained. CNN can diminish the complexity and enhance the efficiency of neural network model training.

Long short-term memory (LSTM) enhances the cyclic neural network by incorporating a cell memory unit, allowing it to store memory over long intervals. It employs gate structures to process information to ensure a strong time correlation between two moments with long time intervals. The flow of LSTM is illustrated in Figure 1.

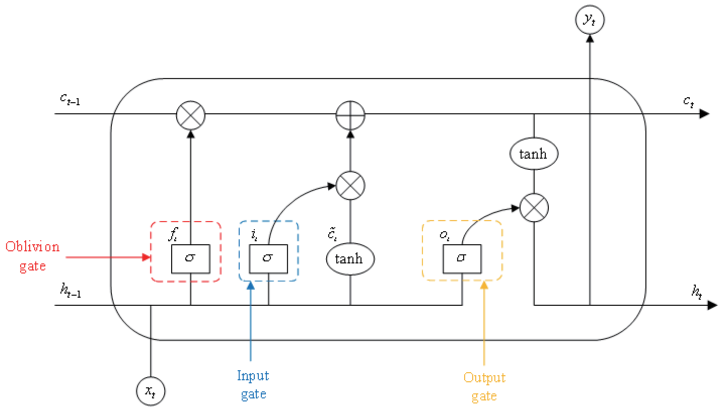

FIGURE 1. Flow for LSTM.

LSTM adds three control gates: oblivion, input, and output. The oblivion gate determines redundant information that should be removed from the unit state based on current information x_t and the output from the previous moment h_{t-1}, which is calculated as follows:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_t), \tag{2.1}$$

where f_t is the activation value for the oblivion gate, σ shows the sigmoid activation function, W_f denotes the weight value, and b_t is the bias.

The input gate updates the memory state for the previous moment c_{t-1} based on x_t and h_{t-1}. The calculation process is as follows:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \tag{2.2}$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c), \tag{2.3}$$

$$c_t = f_t c_{t-1} + i\tilde{c}_t, \tag{2.4}$$

where $\tanh(\cdot)$ defines as the hyperbolic tangent function, c_{t-1} is the memory state for the previous moment, c_t is the memory state for the current moment, b_i and b_c are biases.

The output gate determines the output for the hidden layer at the current moment h_t based on previous information. The formulas used are as follows:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \tag{2.5}$$

$$h_t = o_t \tanh(c_t). \tag{2.6}$$

Bidirectional long short term memory (BiLSTM) is an improved method based on LSTM. It combines forward and reverse LSTM, which takes into account both the current iteration time and reverse timing. It contains all the information of both forward and reverse LSTM, effectively improving the availability of data and the prediction performance of the model. The flow for BiLSTM is shown in Figure 2.

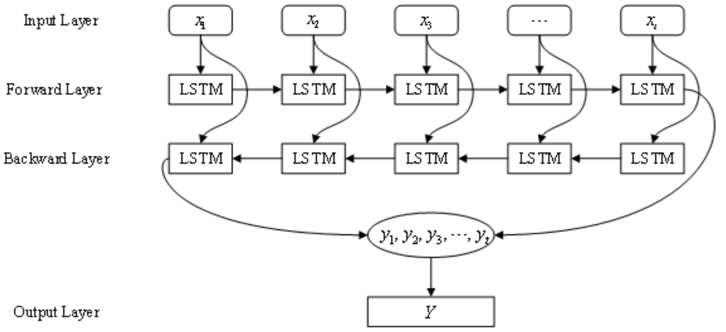

Figure 2. Network structure of BiLSTM.

The mechanism of attention involves ignoring irrelevant information and selecting useful information. Therefore, experts have proposed the attention mechanism. This mechanism assigns different weights to input features based on their influence on the output, thus screening out important features and improving prediction precision. For the output of BiLSTM $h_t = \{h_1, h_2, \cdots, h_r\}$, the attention mechanism is calculated as follows:

$$e_t = \nu \tanh(wh_t + b), \tag{2.7}$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{t} e_j}, \tag{2.8}$$

where e_t is the attention probability distribution at moment t, ν and w are the weight values, b is the bias value.

2.2. **Combination experimental model.**

The proposed model was divided into input, CNN, BiLSTM, attention, and output layers. The data were input into the model through the input layer. The CNN layer extracted the feature of time sequence and generated the feature vector to input into the BiLSTM layer. The BiLSTM and attention layers forecasted the daily newly confirmed cases by learning the rules in the feature vector extracted using CNN, and the prediction results were output through the output layer.

The main parameters used for model training were as follows: the time step set to 5, the optimization algorithm set to Adam, the number of hidden layer nodes set to 16, the training epochs set to 50, and the batch size set to 32. Figure 3 shows the flow of the CNN-BiLSTM-Attention combination model.

Table 1 displays the structure of each layer of the experimental model.

(1)The CNN layer grabbed the features of the data from the input layer, mined and screened time sequence features, and input them to the BiLSTM layer. The

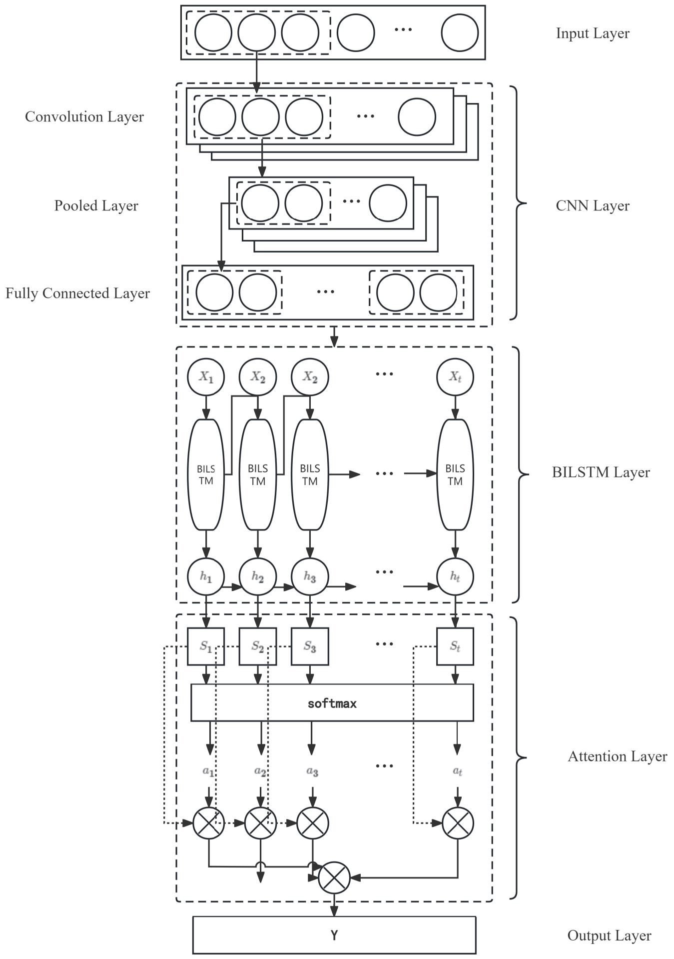

FIGURE 3. Flow diagram of CNN CBiLSTM C Attention composite model.

CNN layer included the convolution pooling, and full connection layers. One-dimensional convolution was selected for the convolution layer based on the characteristics of the COVID-19 time-series data. The convolution kernel size was 16, and tanh was used as the activation function. Then, maxpooling was adopted for pooling on the pooling layer, and the maximum pooling was set to 5. After processing the input vector through the convolution and pooling layers, the dropout layer with a probability of 0.05 was introduced to avoid feature overfitting. Finally, a series of data was output as input to the next BiLSTM layer.

(2)BiLSTM positive and negative structures were used to train the data characteristics of COVID-19 sequences.

(3)The output for each time step for the BiLSTM layer was taken as the input of the attention layer, different weights were added to input features, and they were output to the output layer through the full connection layer.

Table 1. Parameters of CNN CBiLSTM C Attention composite model

Layer number	Name of the layer	Network parameter	Data structure	Activation function
1	Input layer	None	(None, 5, l)	None
2	Convolution layer	Fters = 16, kernel size = 1	(None, 5, 16)	Tanh
3	Pooled layer	pool size = 5	(None, 1, 16)	None
4	Dropout layer	Dropout rate = 0.05	(None, 1, 16)	None
5	BiLSTM layer	Number of nodes: 32	(None, 32)	Tanh
6	Attention layer	Number of nodes: 32	(None, 32)	Sigmoid
7	Fully connected layer	Number of nodes: 1	(None, 1)	Tanh

2.3. **Model performance evaluation index.**

In this study, the predictive performance of the model was evaluated using two indexes: mean absolute error (MAE) and root mean square error(RMSE). The evaluation formula used was as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|y_i^{'} - y_i|^2}, \tag{2.9}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i^{'} - y_i|, \tag{2.10}$$

where i denotes the label of sample data, y_i is the true value of the test set of newly confirmed cases, n represents the total number of sample data, and $y_i^{'}$ is the predicted value for the test set of newly confirmed cases.

3. Results

3.1. **Sample acquisition.**

This study compiled Tencent real-time epidemic data (https://api.inews.qq.com) on COVID-19 from various European countries by Jupyter Notebook (Anaconda3). By checking and comparing the COVID-19 data reports on different Web sites, it was found that the real-time COVID-19 data published by Tencent were relatively comprehensive, with fewer missing values and more periods. Therefore, this study consolidated the COVID-19 data from the 43 European countries and 1 region to obtain the total COVID-19 epidemic data of the entire Europe spanning from March 13, 2020, to November 11, 2021.

3.2. **Normalized processing.**

The epidemic values were extremely large and the calculations were complicated. Therefore, the values were normalized to keep the epidemic data between [0,1] using the following equation:

$$\bar{X}_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}, \tag{3.1}$$

where i is the number of samples, whose value is 1 to T, T denotes the total number of the samples, $X_{\max}$ and $X_{\min}$ are the maximum and minimum values for the samples, respectively.

3.3. **Data denoising using EMD.**

The dataset was divided in the ratio of 8:2. That is, the first 80% of data were used as the model training set, and the remaining 20% of the data were used as the model test set. The predictive performance of the model was assessed using the evaluation indexes RMSE and MAE.

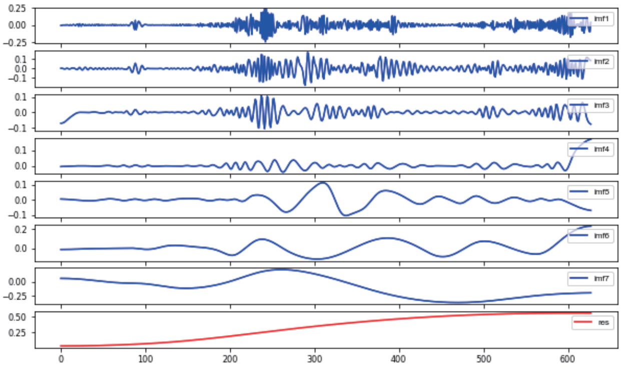

FIGURE 4. EMD decomposition result.

The COVID-19 data were decomposed using EMD. As shown in Figure 4, the original data were decomposed into seven intrinsic mode functions (IMFs) and one residual component (Res). IMFs reflected the distribution of data in different influencing factors and scales. The noise gradually decreased and the data became stable.

3.4. **Analysis of results.**

The trend comparison between the predicted and real values of the original data under eight prediction models is shown in Figure 5.

The predicted effect of eight models is presented in Table 2.

As shown in Figure 5 and Table 2, the predicted trend of the model with EMD decomposition algorithm was better than the actual trend. Both RMSE and MAE were significantly reduced. Upon integrating the EMD decomposition algorithm into various models, the EMD-BiLSTM model exhibited varying degrees of improvement in prediction performance. Compared with BiLSTM model, RMSE and MAE decreased by 27.82% and 29.07%, respectively, with the largest reduction range. The

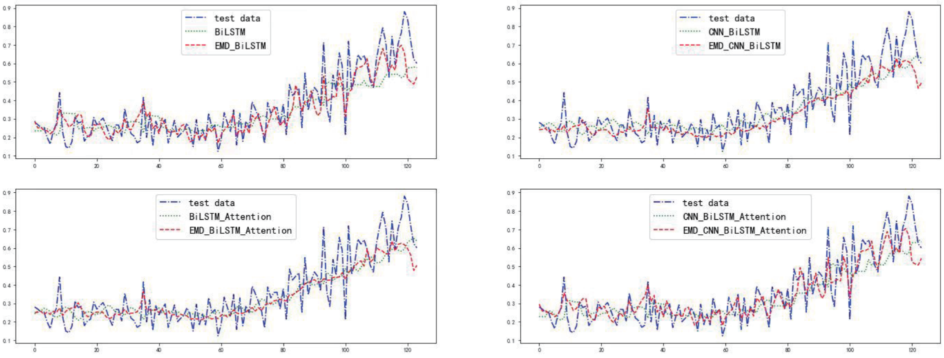

Figure 5. Trend comparison of forecast models.

Table 2. Predicted performance of eight models

Model	RMSE	MAE
BiLSTM	0.1136	0.0853
EMD-BiLSTM	0.0820	0.0605
CNN-BiLSTM	0.1053	0.0818
EMD-CNN-BiLSTM	0.955	0.0733
BiLSTM-Attention	0.1041	0.0815
EMD-BiLSTM-Attention	0.0916	0.0703
CNN-BiLSTM-Attention	0.0982	0.0743
EMD-CNN-BiLSTM-Attention	0.0798	0.0592

RMSE and MAE for the EMD-CNN-BiLSTM model decreased by 9.31% and 10.39% compared with those for the CNN-BiLSTM model. Compared with the BiLSTM-Attention model, RMSE and MAE decreased by 12.01% and 13.74%, respectively. Compared with the CNN-BiLSTM-Attention model, RMSE and MAE decreased by 18.74% and 20.32%, respectively. Compared with the four models without EMD decomposition algorithm, the CNN-BiLSTM-Attention combined model displayed the best prediction performance, and the prediction error was noticeably smaller than that in the other three models. Comparison among the four models using the EMD decomposition algorithm revealed that the EMD-CNN-BiLSTM-Attention combined model demonstrated the most effective prediction performance.

4. Conclusions

This study integrated the benefits of deep learning and neural network prediction algorithms by establishing a combined prediction model based on the bidirectional LSTM neural network. The model incorporated BiLSTM along with CNN and Attention mechanisms. Additionally, several comparison models were constructed to verify the superiority of the experimental model. For the experimental data in the model training without the EMD algorithm, the prediction effect of the CNN-BiLSTM-Attention model was better than that of the other models. In addition, when the experimental data and logarithmic control data were decomposed using the EMD algorithm, it significantly improved the prediction performance of the model and reduced the prediction error. The model proposed in this study displayed better performance in predicting the time series for confirmed COVID-19 cases in Europe, with an RMSE of 0.0798 and MAE of 0.0592. The proposed model achieved higher prediction accuracy compared with the comparison models. It provided a promising method for predicting COVID-19.

References

[1] M. Arfan, K. Shah, T. Abdeljawad, N. Mlaiki and A. Ullah, *A capto power law model predicting the spread of the COVID-19 outbreak in Pakistan*, Alexandria Engineering Journal **60** (2021), 447–456.

[2] V. Arvind, J. S. Kim, B. H. Cho, E. Geng and S. K. Cho, *Development of a machine learning algorithm to predict intubation among hospitalized patients with COVID-19*, Journal of Critical Care **62** (2021), 25–30.

[3] D. Benvenuto, M. Giovanetti, L. Vassallo, S. Angeletti and M. Ciccozzi, *Application of the ARIMA model on the COVID-2019 epidemic dataset*, Data in brief **29** (2020): 105340.

[4] Z. Chaudhry, M. Shawe-Taylor, T. Rampling, T. Cutfield T, G. Bidwell, X. H. S. Chan, A. Last, B. Williams, S. Logan, M. Marks and H. Esmail, *Short durations of corticosteroids for hospitalised covid-19 patients are associated with a high readmission rate*, Journal of Infection **82** (2021), 276–316.

[5] K. Choi and S. B. Thacker, *Mortality during influenza epidemics in the United States, 1967-1978*, American Journal of Public Health **72** (1982), 1280–1283.

[6] S. Farah, Z. Aneela and M. Muhammad, *Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM*, Chaos, Solitons and Fractals **140** (2020): 110212.

[7] S. Hochreiter and J. Schmidhuber, *Long short-term memory*, Neural Computation **9** (1997), 1735–1780.

[8] E. Huang, Z. Shen Z, R. Long, C. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung and H. Liu, *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **454** (1998), 903–995.

[9] M. Jahangir, A. Muheem and M. Rizvi, *Coronavirus (COVID-19): history, current knowledge and pipeline medications*, Journal of Pharmacy and Pharmacology **4** (2020), 1–5.

[10] F. W. Jasper, S. F. Yuan, K. H. Kok, K. To, H. Chu, J. Yang, F. Xing, J. Liu, C. C. Yip, R. W. Poon, H. Tsoi, S. K. Lo, K. Chan, V. K. Poon, W. Chan, J. D. Ip, J.Cai, V. C.Cheng, H. Chen, C. K. Hui and K. Yuen, *A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster*, The Lancet **395** (2020), 514–523.

[11] H. Li, E. Zheng, Z. Zhong, C. Xu, N. Roma, S. Lamkin, S. Lamkin, T. T. Von Visger, Y. Chang and W. Xu, *Stress prediction using micro-EMA and machine learning during COVID-19 social isolation*, Smart Health **106** (2022): 100242.

[12] X. Liu and S. Fong, *Towards a realistic model for simulating spread of infectious COVID-19 disease*, in: Proceedings of the 2020 the 4th International Conference on Big Data and Internet of Things, Association forComputing Machinery, 2020, pp. 96–101.

[13] R. Lucas and S. Gecynalda, *Forecasting daily Covid-19 cases in the world with a hybrid ARIMA and neural network model*, Applied Soft Computing **126**(2022): 109315. doi: 10.1016/j.asoc.2022.109315

[14] J. Luo, Z. Zhang, F. Yao and F. Rao. *Time series prediction of COVID-19 transmission in America using LSTM and XGBoost algorithms*, Results in Physics **27** (2021):104462.

[15] K. L. P. Nguyen, H. Chen, C. Yang and E. Kristiani, *Implementation of a respiratory disease forecasting model using LSTM for central Taiwan*, Information Science and Applications **621** (2020), 441–450.

[16] N. H. Nguyen and I. T. Nabney. *Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models*, Energy **35** (2010), 3674–3685.

[17] D. Parbat and M. Chakraborty, *A python based support vector regression model for prediction of COVID19 cases in India*, Chaos, Solitons and Fractals **138** (2020): 109942.

[18] D. Radojičić and S. Kredatus, *The impact of stock market price fourier transform analysis on the gated recurrent unit classifier model*, Expert Systems with Applications **159** (2020): 113565.

[19] I. Rahimi, F. Chen and A. H. Gandomi, *A review on COVID-19 forecasting models*, Neural Computing and Applications **35** (2023), 23671–23681.

[20] R. V. Shashank, N. B. Sai, A. S. Eric and A. Amod, *Prediction of the number of COVID-19 confirmed cases based on K-means-LSTM*, Array **11** (2021): 100085.

[21] N. S. Siti, K. P. Mohamad and S. A. S. Muhamad, *A network analysis and support vector regression approaches for visualising and predicting the COVID-19 outbreak in Malaysia*, Healthcare Analytics **2** (2022): 100080.

[22] F. Wen, J. Xiao, Z. He and X. Gong, *Stock price prediction based on SSA and SVM*, Procedia Computer Science **31** (2014), 625–631.

[23] J. Wu, K. Leung and G. Leung, *Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study*, The Lancet **395** (2020), 689–697.

[24] Q. Yang, J. Wang, H. Ma and X. Wang, *Research on COVID-19 based on ARIMA model*$^{\Delta}$*–taking Hubei, China as an example to see the epidemic in Italy*, Journal of Infection and Public Health **13** (2020), 1415–1418.

[25] C. Zhang, Q. Lan, X. Mi, Z. Zhou, C. Ma and X. Mi, *A denoising method based on the nonlinear relationship between the target variable and input features*, Expert Systems with Applications **218** (2023): 119585.

[26] T. Zhou, Q. Liu, Z. Yang, J. Liao, K. Yang, W. Bai, X. Lu and W. Zhang, *Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019-nCoV*, Journal of Evidence-Based Medicine **13** (2020), 3–7.

Manuscript received February 27, 2024
revised September 12, 2024

S. Ma
school of mathematics and information science, north minzu university, yinchuan, 750021, China
Ningxia key laboratory of intelligent information and big data processing, yinchuan, 750021, China
E-mail address: `sjma@nmu.edu.cn`

X. Bai
school of mathematics and information science, north minzu university, yinchuan, 750021, China
E-mail address: `sxbxy@stu.nmu.edu.cn`

Z. Hao
school of mathematics and information science, north minzu university, yinchuan, 750021, China
E-mail address: `hz1015731678@163.com`

H. Chen
school of mathematics and information science, north minzu university, yinchuan, 750021, China
E-mail address: `292211325@qq.com`