Forecasting COVID-19 Infection Using Encoder-Decoder LSTM and Attention LSTM Algorithms

Khder Alakkari [1], Alhumaima Ali Subhi [2], Hussein Alkattan [3], Ammar Kadi [4], Artem Malinin [4], Irina Potoroko [4], Mostafa Abotaleb [3], El-Sayed M El-kenawy [*5]

[1] Department of Statistics and Programming, Faculty of Economics,
University of Tishreen, Latakia, P.O. Box 2230, Syria
[2]Department of Food and Biotechnology, South Ural State University,
454080 Chelyabinsk
[3] Electronic and Computer Center, University of Diyala, Baqubah MJJ2+R9G, Iraq
[4]Department of System Programming, South Ural State University, 454080 Chelyabinsk, Russia
[5]Department of Communications and Electronics, Delta Higher Institute of Engineering and
Technology, Mansoura 35111, Egypt
Emails: : khderalakkari1990@gmail.com; alhumaimaali@uodiyala.edu.iq;
alkattan.hussein92@gmail.com; ammarka89@gmail.com; artemmalinin3@gmail.com;
irina_potoroko@mail.ru; abotalebmostafa@bk.ru; skenawy@ieee.org

Abstract

The COVID-19 epidemic has in fact placed the whole community in a dire predicament that has led to numerous tragedies, including an economic downturn, political unrest, and job losses. Forecasting and identifying COVID-19 infection cases is crucial for the government at all levels because the pandemic grows exponentially and results in fatalities. Hence, by giving information about the spread of the epidemic, the government can move quickly at multiple levels to establish new policies and modalities in order to minimize the trajectory of the COVID-19 pandemic's effects on both public health and the economic sectors. Forecasting models for COVID-19 infection cases in the Ural region in Russia were developed using two deep Long Short-Term Memory (LSTM) learning-based approaches namely Encoder–Decoder LSTM and Attention LSTM algorithms. The models were evaluated based on five standard performance evaluation metrics which include Mean Square Error (MSE), Mean Absolute Error (MAE), Root MSE (RMSE), Relative RMSE (RRMSE), and coefficient of determination (R2). However, the Encoder–Decoder LSTM deep learning-based forecasting model achieved the best performance results (MSE=32794.09, MAE=168.56, RMSE=181.09, RRMSE=13.46, and R2=0.87) compared to the model developed with Attention LSTM models.

Keywords: COVID-19; LSTM; RMSE;

1. Introduction

Coronaviruses are a polymorphic group of respiratory viruses that cause acute inflammatory diseases in domestic and farm animals [1]. In humans, the infection, until recently, was observed mainly in the autumn-winter period and was characterized by a mild course [2]. The situation changed dramatically in 2003, when an outbreak of atypical pneumonia caused by the pathogenic SARS-CoV was registered in China. 10 years later, a new outbreak of coronavirus emerged in the form of Middle East respiratory syndrome (MERS-CoV) [3]. The emergence of SARS-CoV2-related illnesses in December 2019 will go down in history as an international emergency that quickly developed into a pandemic in the first few months of 2020 [4]. A new coronavirus, not only new, from the point of view of its molecular and biological features, but in the context of possible difficulties in diagnosis and treatment, features

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

of the clinical course, high risk of critical conditions and complications, high mortality [5]. The disease often leads to severe bronchopulmonary lesions, ranging from a dry debilitating cough to acute respiratory distress syndrome [6].

Coronavirus infection worsens the operative memory of most patients over the age of 25, but a year after the acute phase of COVID-19, this function is fully restored, a study by British scientists has shown [7]. At the same time, the harder a person suffers from the disease, the more his memory suffers (long-term and short-term covid) [8]. One of the most crucial elements in determining the mortality and morbidity linked to COVID-19 was how patients who required critical care were treated [9]. It is a significant issue for healthcare systems worldwide to prescribe COVID-19 medication to patients who require immediate or urgent respiratory care [10].

Machine learning to simulate cases of COVID-19 infection using an encoder-decoder and attention based on a deep regression model of long-term short-term memory is necessary to combat coronavirus infection and stop its global expansion, as well as for the rehabilitation of patients [11]. Intelligent healthcare is increasingly relying on artificial intelligence, notably machine learning algorithms. Machine learning for simulation includes networks that can learn from unlabeled or unstructured data without supervision [12]. COVID-19 applications are software applications that make extensive use of deep learning, using digital tracking to help track contacts in response to the COVID-19 pandemic, that is, the process of identifying individuals who may have been in contact with an infected person in order to prevent the wider spread of the disease. To prevent the spread of infection, three main factors must be taken into account: determining the cause, taking preventive measures and trying to develop an effective treatment [13].

In Russia, as of the end of 2022, there are more than 20 million confirmed cases and 386 thousand deaths. To date, research is continuing aimed at solving the problems associated with this disease, as well as containment mechanisms and public health policies [14]. Quarantine procedures were aimed at slowing down or stopping the spread of COVID-19, in order to improve the efficiency of medical care. In this regard, it is recommended to develop and implement public health strategies [14]. Among the set of machine learning methods based on learning representations, rather than specialized algorithms for specific tasks, are deep learning models that can help in the development of forecasting models [15]. In recurrent neural networks, the connections between the elements take the shape of a directed sequence. An artificial recurrent neural network (RNN) architecture called long-term short-term memory (LSTM) is utilized in deep learning [16]. Although numerous neural networks (NNS) have been reported in the past and have also been able to produce an accurate prediction of what will happen in the future, RNN and LSTM are employed in SARS-CoV-2 prediction because they use transitory data [17].

The Kermak-Mckedrick model (SIR Model) is one of the simplest experimental models in which the dynamics of groups of susceptible, infected and recovered individuals is described using systems of differential equations [18]. The model consists of three "cells". S: the number of people susceptible to infection, that is, those people who are not immune to this virus and can potentially become infected. I: the number of infected at some point in time [19]. These are infected people who can infect susceptible people. R: the number of people who have been ill, have immunity, or the number of deceased persons [20]. That is, these are people who were infected and either recovered from the disease and got into a remote compartment, or died. Such a model can be used to calculate indicators such as the spread of the disease, the total number of infected or the duration of the epidemic, as well as to evaluate various epidemiological parameters, such as the reproductive number [21].

These simulations can demonstrate how different public health policies might influence the course of an epidemic, for instance, how precautions can influence the rate of COVID-19's spread [22]. The disadvantage of the Kermak-Mckedrick model is the lack of flexibility – the inability to take into account changes in parameters such as: new mutations of the virus and strain, restrictive measures, vaccination [23]. These models are based on presumptions that, given the circumstances of the SARS-CoV-2 pandemic, seem to be wrong. Hence, more advanced modeling techniques and in-depth understanding of the biology and epidemiological characteristics of the disease are required in order to predict a pandemic [24]. In addition to more conventional techniques, there are two models (RNNS and LSTM) that can forecast temporal data. Recurrent neural networks (RNNS), a form of artificial neural network built from direct communication networks that exhibits behavior resembling that of the human brain, have been used to handle time series and sequential data [25]. An advanced form of recurrent neural network design that can recognize long-term dependencies is the LSTM. The average projected errors for COVID-19 infection cases using machine learning models are almost on par with statistical model mistakes. Long-term time series can be predicted by machine learning techniques [26].

For NLP works, the Encoder-Decoder long short-term memory (LSTM) was developed. A recurrent neural network is the basis of the Encoder-Decoder architecture (RNN) [27]. When compared to other approaches in the literature, particularly those used for text translation, it performed well [28]. Current applications of the Encoder-Decoder LSTM include the prediction of power consumption [29], metal temperature [30], air pollution [31], behavior [32], and gas concentration [33]. Therefore, modern deep units must be used to create the LSTM core for Encoder-Decoder architecture. In addition, using an Encoder-Decoder architecture to anticipate the spread of a pandemic is a pressing need.

Confirmed cases of Covid-19 from the past are often input into an auto-encoding based architecture in the form of time series data. As a sequential self-learning technique, the provided sequential AE is built from a pair of independent Bi-LSTM based encoder and decoder components. Then, we use the encoding component of the proposed AE architecture to obtain the combined (backward/forward) hidden states of the input sequences, expressed as the imported number of Covid-19 instances during a particular time period [34].

In this study, we use machine learning to simulate COVID-19 infection cases in the Ural region using an encoder-decoder and attention based on a deep regression model of long-term short-term memory.

2. Methodology

A. Long short-term Memory (LSTM)

LSTM networks are an improved model of recurrent neural networks (RNN), first introduced by the two scientists [35][36], the main goal of its development was to avoid the problems of simple RNN and to obtain better results. All RNN contain a series of repetitive patterns, and in traditional RNN [37], these patterns are in the form of a single layer of recurrent neurons as shown in figure 1.

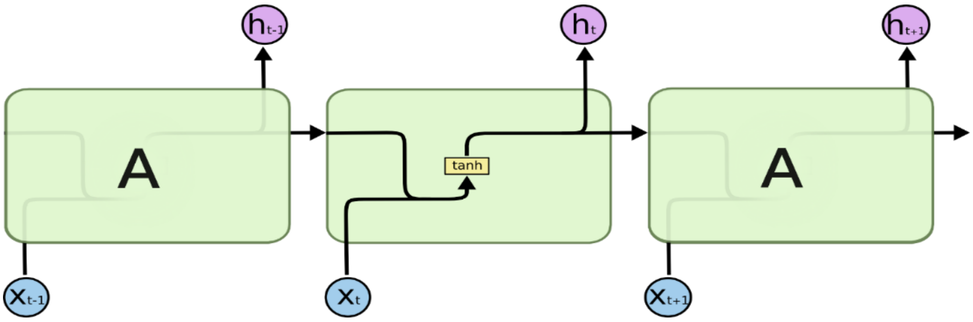

Figure 1: The recurrent form within a simple recurrent network

Figure 1 shows how the neural network takes advantage of the lag information and the lead information of the studied phenomenon. Networks with LSTM also contain a chain, but the shape of this chain is different, as it contains 4 layers instead of 1 layer. Which is shown in figure 2.

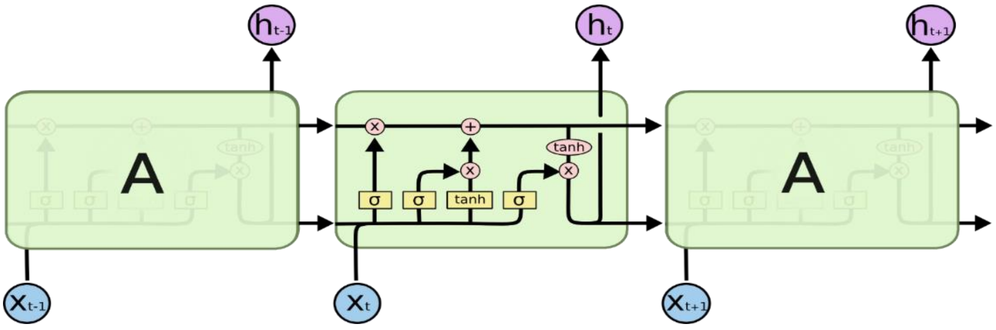

Figure 2: Recurrent form for LSTM model includes 4 layers

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

Figure 2 explains us the mechanism work of LSTM model, the input information is passed to the forget layer, at which point the model decides to: (a) keep the information in the past and use it for prediction, or (b) forget the information and rely on the instantaneous state, then send this information to a tanh function to normalize the information and extract features and patterns and remove noise from them The main goal of designing LSTM is to reduce long-term dependency and its negative impact on the learning process. In addition to the four gates that the network depends on for its work, it helps the network to remember the most important information, which greatly improves the quality of the output [38]. The key to working with LSTM networks is the cell state, as it is considered like a conveyor belt, as it passes along the entire chain and undergoes slight changes during its passage, and therefore it is a good way to keep information unchanged.

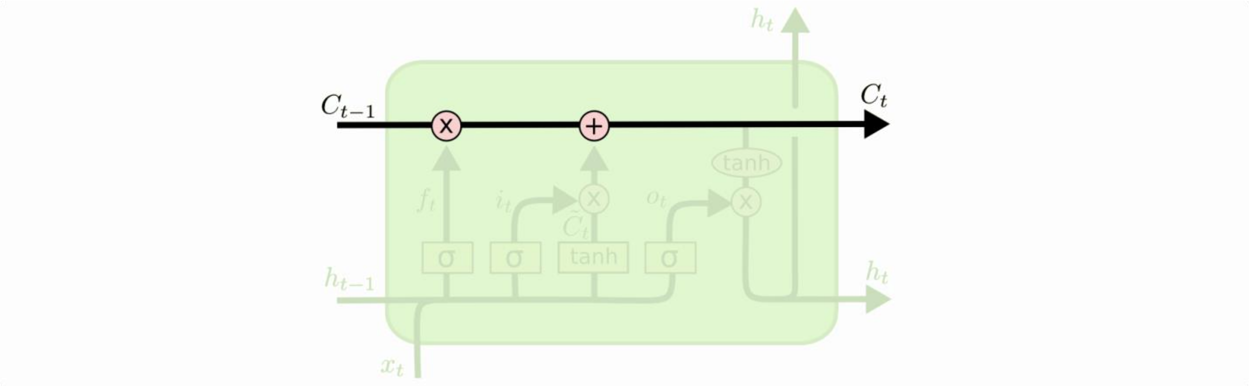

Figure 3: layer state for LSTM model

Figure 3 shows: the state cell within the network, where LSTM networks have the ability to change information within the state cell through an architecture based on logic gates.

These gates consist of a set of neuron layers ending with a sigmoid and a set of positive multiplication operations.

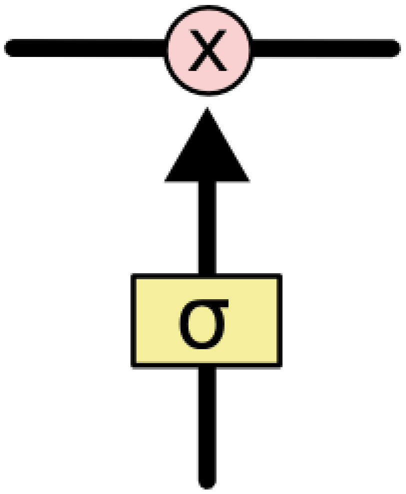

Figure 4: Logic gate within LSTM model

The output of the sigmoid layer is between 0 and 1, and its value specifies the amount of information to be allowed to pass from each cell element. LSTM networks contain three logic gates to control the state of the cell. The LSTM model making process consists of three steps [39]:

First step: A decision is made about what information to keep and what is better to forget from the state cell, and this process takes place within the sinusoidal exponential activation function layer, which is called the forget gate [40], through the following equation:

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

$$f_t = \sigma(W^f x_t + U^f h_{t-1} + b^f) \quad (1)$$

Where f_t is updated value; σ is the sigmoid layer (or nonlinear function); x_t represents a sequence of length t; b is constant bias; h represents RNN memory at time step t; and W and U are weight matrices.

Second step: It is represented by specifying the information to be stored in the state cell and it consists of two parts: first a functional layer called input gate which is responsible for determining the changing value and second a layer that does this ends with the exponential shadow activation function Tanh forming a ray of new candidate values $\boldsymbol{u_t}$. Add it to the status cell, and the next step is to merge the work of the two layers to change the value of the cell status [41]. Which is represented by the following equation:

$$i_t = \sigma(W^i x_t + U^i h_{t-1} + b^i) \quad (2)$$

$$u_t = tanh(W^u x_t + U^u h_{t-1} + b^u) \quad (3)$$

Where i_t is the updated value; u_t is new candidate values; σ is the sigmoid layer (or nonlinear function); x_t is a sequence of length t; b is constant bias; h is RNN memory at time step t; and W and U are weight matrices.

Third steps: Changes the value of the previous state cell, c_{t-1}, to the new value c_t, where we multiply the value of the old state byf_t , then add $i_t u_t$, [42] which is the new value multiplied by the Boost rate resulting from the shadow's exponential activation function:

$$c_t = i_t u_t + f_t c_{t-1} \quad (4)$$

Where c_t represents a memory cell and f_t represents a value between 0 and 1 produced by the forget gate. Specifically, a value of 0 denotes that the value is nullified, whereas a value of 1 indicates that it is retained [40].

Last step: It is supposed to determine the final output and is based on the output of the state cell, but after making some adjustments: First we pass the value on the sinusoidal exponential activation function layer to determine which part of the state cell we select, then we pass the value of the cell state by the exponential activation function of the shadow and multiply it by the output of the layer of the pocket of the exponential activation function [43]:

$$o_t = \sigma(W^o x_t + U^o h_{t-1} + b^o) \quad (5)$$

$$h_t = o_t tanh_t c_{t-1} \quad (6)$$

Where o_t is an output gate and h_t is a value between [1, -1].

B. Encoder – Decoder LSTM

Encoder – Decoder LSTM model were primarily designed to address the sequence-to-sequence problem, which is called seq2seq for short. This problem can be described as the number of sequence elements at the time of input differs from the number at the time of output, which leads to the loss of important information [44]. The modeling problem in this case is that the length of the input sequence may differ from the length of the output sequence due to the multiple lengths of the input and output steps. Accordingly, Encoder – Decoder LSTM is used, which is one of the methods that have proven effective to avoid the problem of seq2seq [45]. This architecture consists of two models: one to read the input sequence and encode it into a fixed-length vector, and a second to decode the fixed-length vector and output the predicted sequence [46][47]. Which can be merged by encoder-decoder LSTM specially designed for seq2seq problems. The main objective of the coding phase is to extract more features and information from the input time series data. The data of an asymmetric sequence of length $X = \{x\_1, x\_2, .. x\_t\}$ is used as input and the encoder encodes the sequence into a fixed length state vector c, which is used as input to the decoder [45]. In the decoder stage, the decoder decodes the state vector c and predicts the next time sequence Y by integrating the input data for the current time.

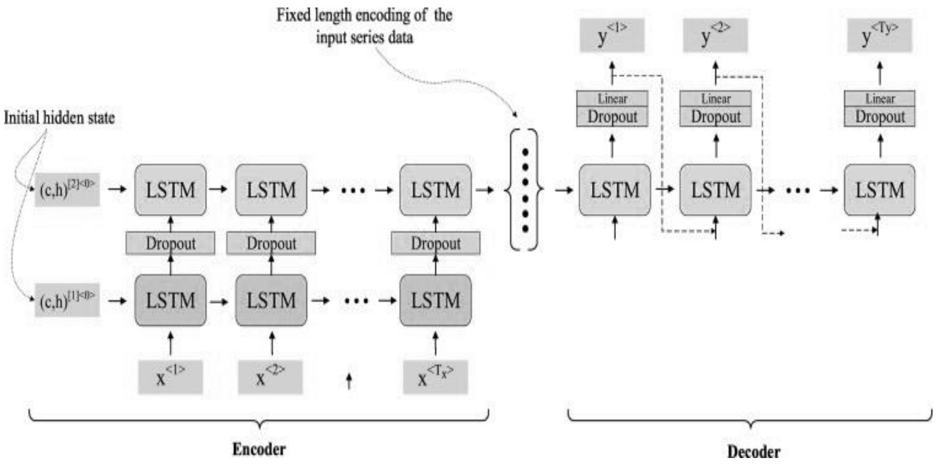

Figure 5: The mechanism of work of the Encoder-Decoder LSTM model

Figure 5 show us the mechanism work of Encoder – Decoder LSTM model, the hidden layer state h is evolved each time the input data is read. When reading the end of the sequence X, the hidden layer variable h, can be thought of as a summary of the input sequence. Which means that the features and information in the sequence have been extracted and mapped in h. For Encoder part The hidden states $h_i$are computed using the formula:

$$h_t = f(W^{hh}h_{t-1} + W^{hx}x_t) \quad (7)$$

With this simple formula only the appropriate weights are applied to the previous hidden state $h_{t-1}$and the input vectorx_t. For decoder part any hidden state $h_{t-1}$is computed using the formula:

$$h_t = f(W^{hh}h_{t-1}) \quad (8)$$

The output $y_t$at time step t is computed using the formula:

$$y_t = softmax\ (W^s h_t) \quad (9)$$

We calculate the exits using the hidden state at the current time step together with the respective weight W(S). Softmax is used to create a probability vector that will help us determine the final output.

C. Attention LSTM

Encoder – Decoder LSTM models are widely used because of their superiority in the fields used. However, with a long sequence of inputs, as in the case of time series, the ED LSTM model encodes a fixed length input sequence [48]. This imposes limits on the length of input sequences that are in the learning phase and causes worse performance for long input sequences [49][50]. Attention is used with the aim of freeing the decoder structure from its internal fixed-length representation. The attention mechanism allows obtaining different information of first-order and lower-order importance and not just the first-order important information. It is described as mapping a query and a set of key-value pairs to an output, where the query, keys, values, and output are all vectors [51]. The output is computed as a weighted sum of the values, where the weight assigned to each value is computed by the query's compatibility function with the corresponding key [52].

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

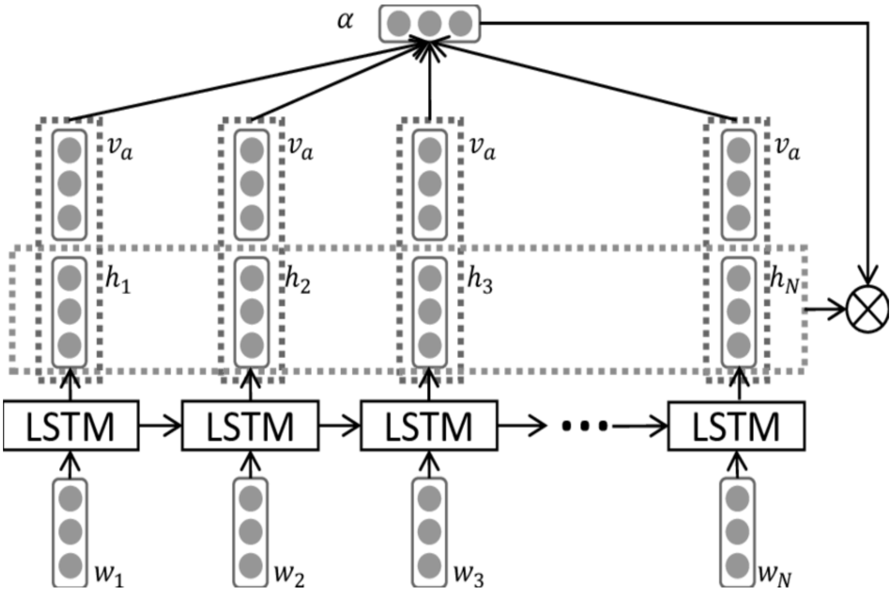

Figure 6: The architecture of attention – based LSTM

Figure 6 show us how attention LSTM work, it includes the first step, we map x_t to h_t :

$$h_t = f_t(h_{t-1}, x_t) \qquad (10)$$

Where f is nonlinear activation function, $h_t \in R_t^s$ is hidden state at time t, s is size of hidden state, in the second step, an attention mechanism is built through a stochastic attention model. For a particular feature sequence $x^k = (x_1^k, \dots, x_m^k)$. From the previous hidden state h_{t-1} and cell state c_{t-1} in the LSTM unit, it is determined [53]:

$$a_t^k = V^T tanh\,(W_1.[h_{t-1}, c_{t-1}] + W_2 x^k) \qquad (11)$$

$$\beta_t^k = softmax(a_t^k) = \frac{exp\ (a_t^k)}{\sum_{i=1}^{n} exp\ (a_t^k)} \qquad (12)$$

Where$V$: is vector, $W1$ and $W2$ are matrices and both learnable parameters by model. a_t^k: is vector has length m and its $h_t - i$ measures the importance of $h_k - i$ input features sequence at time t. and normalized by softmax. β_t^k: is an attention weight, which contains a score of how much attention should be put on $h_k - I$ feature sequences.

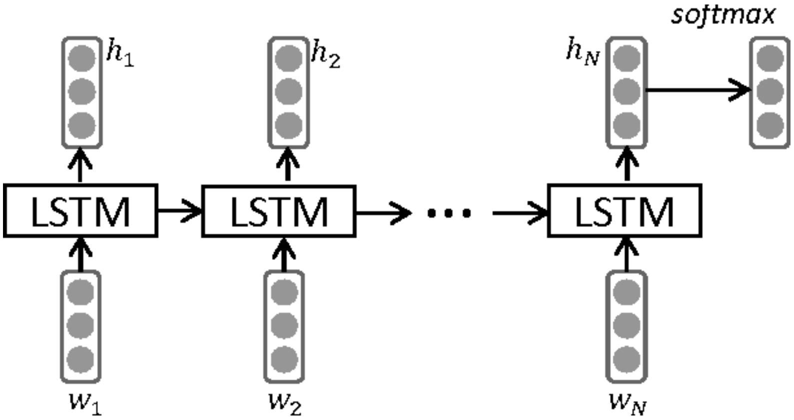

Figure 7: The architecture of attention – based LSTM Normalized by softmax

From Figure 7, the different information in the sequence length is sent to the SoftMax function to calibrate to uniform weights. Accordingly, the output of the attention model at time t of weighted input feature u_t is as follows:

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

$$u_t = (\beta_t^1 x_t^1, \beta_t^2 x_t^2, \dots, \beta_t^n x_t^n)^T \tag{13}$$

Thus, the x in equation 1 is replaced by the weights u_t in the current equation to develop the attention model. It is possible to obtain attention-based time series with better features than input sequence elements.

3. Performance indicator

We use indicators to evaluate the performance of the models used to determine their ability to explain the features and information contained in the data. This is done by examining the extent to which the estimated values using the model correspond to the actual values, taking into account the avoidance of an under fitting problem that may appear from the training data, and an over fitting problem that appears through the test data. The following performance indicators include:

- Mean Square Error (MSE):

$$\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n} \tag{14}$$

- Mean Absolute Error (MAE):

$$\frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_t\right| \tag{15}$$

- R-Squared:

$$\frac{Cov(y_t, \hat{y}_t)}{\sqrt{V(y_t)\,V(\hat{y}_t)}} \tag{16}$$

- Root Mean Square Error (RMSE):

$$\sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}} \tag{17}$$

- Relative Root Mean Square Error (RRMSE):

$$\sqrt{\frac{\frac{1}{n}\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{\sum_{t=1}^{n}\left(\hat{y}_t\right)^2}} \tag{18}$$

Where $\hat{\mathbf{y}}_{\mathbf{t}}$ the forecast is value; $\mathbf{y}_{\mathbf{t}}$ is the actual value; and **n** is the number of fitted observed. The smaller the values of these indicators, the better the performance of the model. Table 1 show us the value of the performance indicators of the test data for SARS-CoV-2 infection cases in Ural region. The results show superiority of the (encoder-decoder) LSTM models as imposing restrictions on the length of the input sequence gives better performance than varying the sequence length in the SARS-CoV-2 infection cases data.

Table 1: Comparison of SARS-CoV-2 modelling evaluation for testing dataset (10%)

Model	MSE	MAE	R-Squared	RMSE	RRMSE
(encoder-decoder) LSTM	32794.09	168.56	0.87	181.09	13.46
Attention LSTM	55844.76	226.41	0.77	236.32	15.37

We found from the table lower values for (MSE – MAE – RMSE – RRMSE) and this is evidence that the model data estimated from the actual data are closer in the testing phase. We also found a greater value for the coefficient of determination (R Squared), which indicates the ability of the model to capture variations in the number of SARS-CoV-2 infection cases in Ural region. Table 2 shows the most important descriptive statistics of SARS-CoV-2 infection cases in Ural region.

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

Table 2 : Descriptive statistics of SARS-CoV-2 infection cases in Ural region

Mean	1679.02
Standard Error	89.90
Median	913.00
Mode	0.00
Standard Deviation	2896.24
Sample Variance	8388220.17
Kurtosis	20.54
Skewness	4.35
Range	20539.00
Minimum	0.00
Maximum	20539.00
Sum	1742825.00
Count	1038.00
Confidence Level (95.0%)	176.40

We note from the table 2 that the values of the mean, median, and mode differ, and this indicates that the data is not distributed according to a normal distribution, and therefore it is not possible to rely on the mean and standard deviation in interpretation because it is not robust to outliers. It is possible to rely on median that indicates that the daily rate of infections swallowed 44 cases during the daily period 12/3/2020 to 13/1/2023. The standard error in the table indicates a low value (89.9) indicating that the sample is well representative of the study population. The large size of sample variance (8388220.17) indicates significant changes in the number of SARS-CoV-2 infection cases in Ural region. As the number of infection cases developed from its minimum value (0) on 12/3/2020 to its maximum value (20539) on 12/2/2022. The value of kurtosis (20.54) indicates the Leptokurtic of the data distribution, that is, the presence of values that are far from the median. A positive skewness value (4.35) indicates that the distribution is skewed to the right, meaning that frequencies whose values are greater than the median are more than those whose values are smaller. It is expected, at a confidence level 95%, that the prediction of the number of SARS-CoV-2 infection cases in Ural region will be within the range (913±176.4). Figure 8 shows us the actual and estimated data using the (encoder-decoder) LSTM model in the training phase.

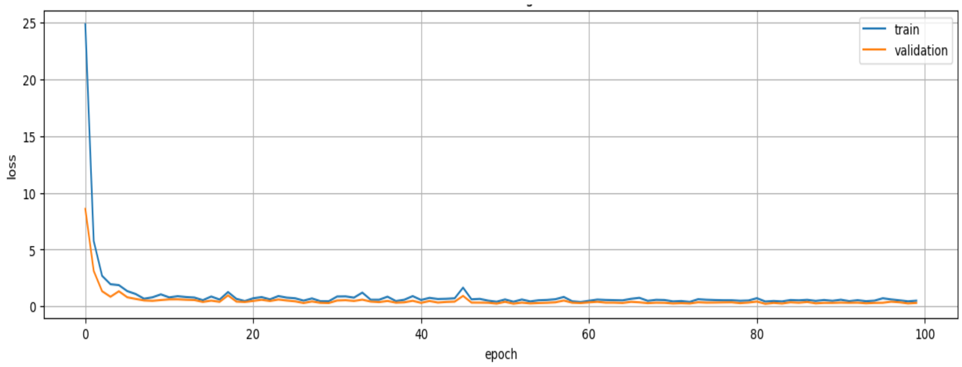

Figure 8: Model train vs validation loss using (encoder-decoder) LSTM

At the loss of validation, we use the model to find a convergence between the actual and estimated values, while the model is able to capture discrepancies in the actual data. By mapping, we rule out the existence of an underfitting problem in the estimate. Figure 9 show us actual and predicted value for COVID19 infection cases in ural region in testing phase, where we find a convergence between

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

the actual values and the expected values, and the return of the number of infections to low levels after the spread of the vaccine in Ural regions. Figure 9 shows the actual and estimated data more clearly.

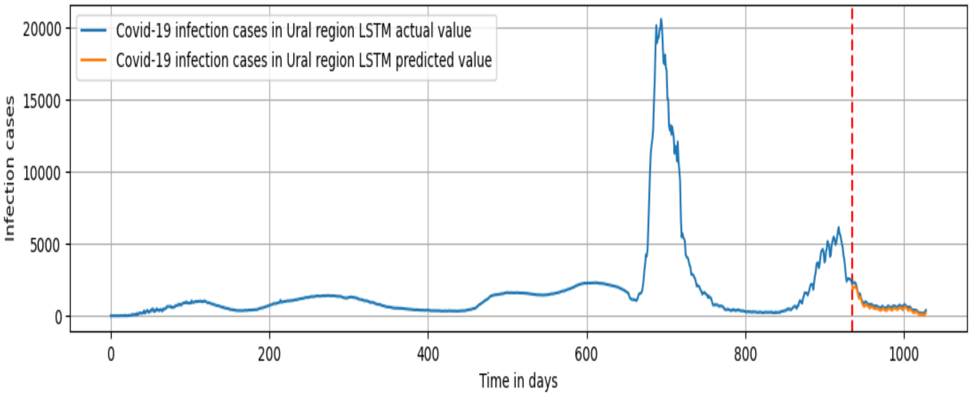

Figure 9: COVID-19 infection cases in Ural region (encoder- decoder) LSTM

Figure show us actual and predicted value for COVID19 infection cases in ural region in testing phase, where we find a convergence between the actual values and the expected values, and the return of the number of infections to low levels after the spread of the vaccine in Ural regions. The following figure shows the actual and estimated data more clearly.

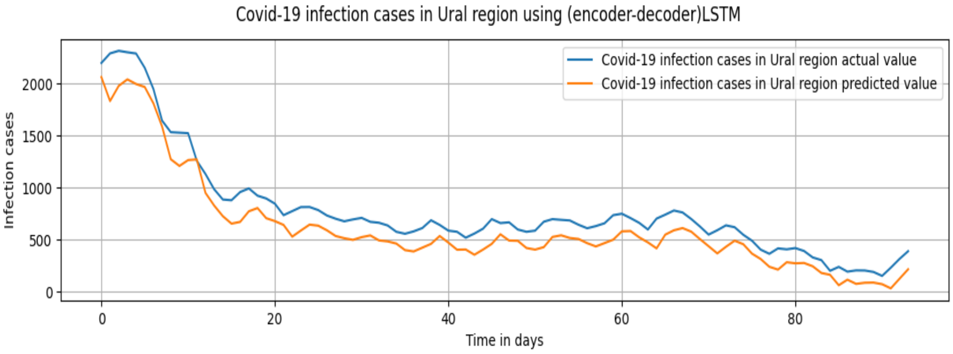

Figure 10: COVID-19 infection cases in Ural region using (encoder- decoder) LSTM

We observe from the figure 10 by capturing the estimation data using the (encoder-decoder) LSTM model for actual data variances and rule out an over fitting problem. Figure 11 show us actual and estimated data using Attention LSTM model on training phase.

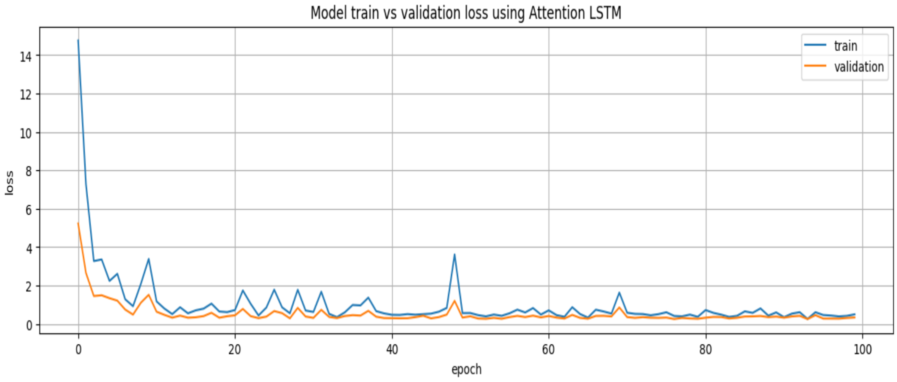

Figure 11: COVID-19 infection cases in Ural region using attention (encoder-decoder) LSTM

We note that the estimation data does not capture changes in the actual data, and here, through the graph, we expect that there is an under fitting problem in the model's estimations, and therefore these

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

results cannot be adopted. Figure 12 show us actual and estimated data using Attention LSTM model in testing phase

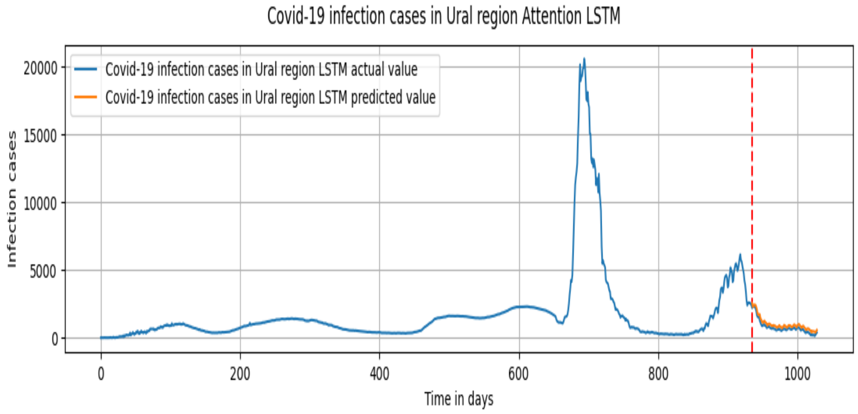

Figure 12: COVID-19 infection cases in Ural region using attention LSTM

Figure 13 show us actual and estimated data using Attention LSTM model in testing phase, The following figure shows these estimates more clearly:

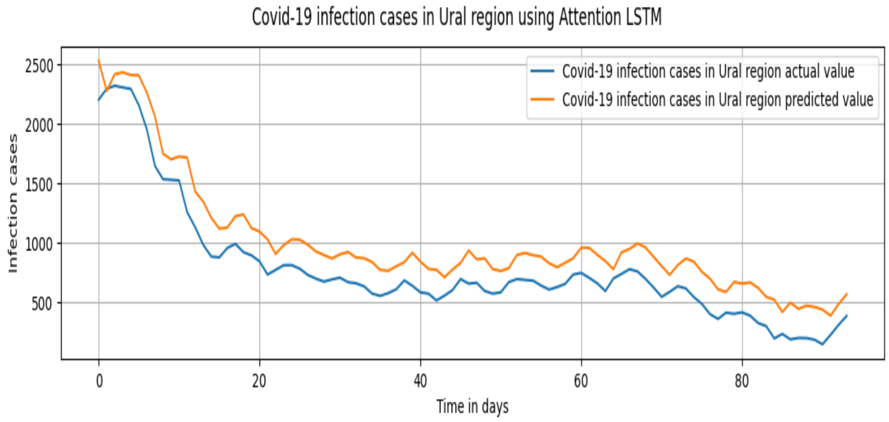

Figure 13: COVID-19 infection cases in Ural region using attention LSTM

Figure 13 show us that the estimation data using the Attention LSTM model has higher variances than the variances of the actual data, and thus we conclude that there is an over fitting problem in the model estimations. Therefore, the estimates of this model cannot be supported.

4. Conclusion

As is well known, the epidemic has an impact on all countries in the world. This research paper examined the role of some deep learning approaches in assisting governmental and medical organizations. In this work, we compared two learning-based Deep Long Short-Term Memory (LSTM) approaches, namely the Encoder Decoder LSTM and Attention LSTM algorithms, to predict COVID-19 infection cases in the Ural region of Russia. The learning models were assessed based on the five popular performance assessment standards, including MSE, MAE, RMSE, RRMSE and R2. However, the deep learning predictive models based on Encoder Decoder LSTM achieved the best performance results compared to the model developed with the Attention LSTM. In order to understand, analyze and collect the latest developments in this field of research, this type of study should be conducted in the future. It can be useful for policy makers and future researchers.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] Cleary, S. J., Pitchford, S. C., Amison, R. T., Carrington, R., Robaina Cabrera, C. L., Magnen, M., ... & Page, C. P., Animal models of mechanisms of SARS-CoV-2 infection and COVID-19 pathology. British journal of pharmacology, 177(21), 4851-4865, 2020.

[2] Sachs J., Schmidt-Traub G., Kroll C., Lafortune G., Fuller G., Woelm F., Sustainable Development Report 2020: The Sustainable Development Goals and COVID-19 Includes the SDG Index and Dashboards, Cambridge University Press: Cambridge, UK, 2021.

[3] Shekerdemian L.S., Mahmood N.R., Wolfe K.K., Riggs B.J., Ross C.E., McKiernan C.A., Heidemann S.M., Kleinman L.C., Sen A.I., Hall M.W., et al., Characteristics and Outcomes of Children with Coronavirus Disease (COVID-19) Infection Admitted to US and Canadian Pediatric Intensive Care Units. JAMA Pediatr. , 174, 868–873, 2020.

[4] Upadhyay S. K., Singh R., Singh M., Kumar V., Yadav M., Aggarwal D., Sehrawat N., COVID-19 in republic of India: A report on situation and precautionary strategies to global pandemic. Bull Environ Pharmacol Life Sci, 9(6), 39-48, 2020.

[5] Koliaki C., Tentolouris A., Eleftheriadou I., Melidonis A., Dimitriadis G., Tentolouris N., Clinical management of diabetes mellitus in the era of COVID-19: practical issues, peculiarities and concerns. Journal of clinical medicine, 9(7), 2288, 2020.

[6] House, C., Naseefa N., Palissery S., Sebastian H, Corona viruses: A review on SARS, MERS and COVID-19. Microbiol. Insights, 14, 2021.

[7] Kovoor J. G., Scott N. A., Tivey D. R., Babidge W. J., Scott D. A., Beavis V. S., Frydenberg M. , Proposed delay for safe surgery after COVID-19. ANZ Journal of Surgery, 91(4), 495-506, 2021.

[8] de Palma, A., & Vosough, S. (2021). Long, medium, and short-term effects of COVID-19 on mobility and lifestyle. CY Cergy Paris Université, cnrs.

[9] Elhadi M., Alsoufi A., Abusalama A., Alkaseek A., Abdeewi S., Yahya M., ... & Msherghi, A. Epidemiology, outcomes, and utilization of intensive care unit resources for critically ill COVID-19 patients in Libya: A prospective multi-center cohort study. Plos one, 16(4), 2021.

[10] Pierce J., & Stevens M. P., COVID-19 and antimicrobial stewardship: lessons learned, best practices, and future implications. International Journal of Infectious Diseases, 113, 103-108, 2021.

[11] Fehaid Alqahtani, Mostafa Abotaleb, Ammar Kadi, Tatiana Makarovskikh, Irina Potoroko, Khder Alakkari, Amr Badr. Hybrid Deep Learning Algorithm for Forecasting SARS-CoV-2 Daily Infections and Death Cases. Axioms, 11(620), 1 – 19, 2022.

[12] Sarker I. H., Machine learning: Algorithms, real-world applications and research directions. SN computer science, 2(3), 160, 2021.

[13] Mbunge E., Akinnuwesi B., Fashoto S. G., Metfula A. S., Mashwama P., A critical review of emerging technologies for tackling COVID-19 pandemic. Human behavior and emerging technologies, 3(1), 25-39, 2021.

[14] Kareem, F. Q., & Abdulazeez, A. M.. Ultrasound medical images classification based on deep learning algorithms: a review. Fusion: Practice and Applications, 3(1), 29-42, 2021 https://doi.org/10.54216/FPA.030102.

[15] Abdulrahma, S. A., & Salem, A. B. M. An efficient deep belief network for Detection of Coronavirus Disease COVID-19. Fusion: Practice and Applications, 2(1), 5-13, 2020 https://doi.org/10.54216/FPA.020102

[16] Pal S., Ghosh S., Nag A. Sentiment analysis in the light of LSTM recurrent neural networks. International Journal of Synthetic Emotions (IJSE), 9(1), 33-39, 2018.

[17] Tanıma Ö., Al-Dulaimi A., Harman A.G.G. Estimating and Analyzing the Spread of COVID-19 in Turkey Using Long Short-Term Memory, In Proceedings of the 2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 17–26, 2021.

[18] Chen Y., Liu F., Yu Q., Li T. Review of fractional epidemic models, Applied mathematical modelling, 97, 281-307, 2021.

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

[19] Sadria M., Layton A. T., Modeling within-host SARS-CoV-2 infection dynamics and potential treatments. Viruses, 13(6), 1141, 2021.

[20] Netea M. G., Giamarellos-Bourboulis E. J., Domínguez-Andrés J., Curtis N., van Crevel R., van de Veerdonk F. L., Bonten M.,Trained immunity: a tool for reducing susceptibility to and the severity of SARS-CoV-2 infection. Cell, 181(5), 969-977, 2020.

[21] Kaplan E. H., Wang D., Wang M., Malik A. A., Zulli A., Peccia J., Aligning SARS-CoV-2 indicators via an epidemic model: application to hospital admissions and RNA detection in sewage sludge, Health care management science, 24, 320-329, 2021.

[22] Baral S. D., Mishra S., Diouf D., Phanuphak N., Dowdy D., The public health response to COVID-19: balancing precaution and unintended consequences. Annals of epidemiology, 46, 12, 2020.

[23] Chen K., Pun C. S., Wong H. Y., Efficient social distancing during the COVID-19 pandemic: Integrating economic and public health considerations. European journal of operational research, 304(1), 84-98, 2023.

[24] Maziarz M., Zach M., Agent-based modelling for SARS-CoV-2 epidemic prediction and intervention assessment: A methodological appraisal. Journal of Evaluation in Clinical Practice, 26(5), 1352-1360, 2020.

[25] Agarwal A., Mishra A., Sharma P., Jain S., Ranjan S., Manchanda R., Using LSTM for the Prediction of Disruption in ADITYA Tokamak, arXiv 2020, preprint. arXiv:2007.06230.

[26] Abotaleb M.S., Makarovskikh T., Analysis of Neural Network and Statistical Models Used for Forecasting of a Disease Infection Cases. In Proceedings of the 2021 International Conference on Information Technology and Nanotechnology (ITNT), Samara, Russia,1-7, 20–24 September 2021.

[27] Shahin A. I., Almotairi S., A deep learning BiLSTM encoding-decoding model for COVID-19 pandemic spread forecasting, Fractal and Fractional, 5(4), 175, 2021.

[28] Wang T., Chen P., Rochford J., Qiang J., Text simplification using neural machine translation. In Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 30, 12–17, 2016.

[29] Du S., Li T., Yang Y., Horng S.J., Multivariate time series forecasting via attention-based encoder–decoder framework. Neurocomputing, 388, 269–279, 2020.

[30] Laubscher R., Time-series forecasting of coal-fired power plant reheater metal temperatures using encoder-decoder recurrent neural networks. Energy, 189, 2019.

[31] Zhang B., Zou G., Qin D., Lu, Y., Jin Y., Wang H., A novel Encoder–Decoder model based on read-first LSTM for air pollutant prediction. Sci. Total Environ. , 765, 144507, 2021.

[32] Zerkouk M., Chikhaoui B., Spatio-temporal abnormal behavior prediction in elderly persons using deep learning models, Sensors, 20, 2359, 2020.

[33] Lyu P., Chen N., Mao S., Li M., LSTM based encoder-decoder for short-term predictions of gas concentration using multi-sensor fusion. Process. Saf. Environ. Prot. , 137, 93–105, 2020.

[34] Pham P., Pedrycz W., Vo B., Dual attention-based sequential auto-encoder for Covid-19 outbreak forecasting: A case study in Vietnam. Expert Systems with Applications, 203, 117514, 2022.

[35] Hochreiter S., Hochreiter J., Long short-term memory. Neural computation, 8(9), 1735-1780, 1977.

[36] Gers F., Eck D., Schmidhuber J., Applying LSTM to time series predictable through time-window approaches. Neural Nets WIRN Vietri-01, 193-200, 2002.

[37] Huynh H., Dang L., Duong D., A new model for stock price movements prediction using deep neural network. In Proceedings of the Eighth International Symposium on Information and Communication Technology, 57-62, 2017.

[38] Zhang X., Liang X., Zhiyuli A., Zhang S., Xu R., Cheng Z., & et al., AT-LSTM: An attention-based LSTM model for financial time series prediction. In IOP Conference Series: Materials Science and Engineering, 569 (5), 052037, 2019.

[39] Song X., Liu Y., Xue L., Wang J., Zhang J., Wang J., et al., Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. Journal of Petroleum Science and Engineering, 186, 2020.

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023

[40] Van Houdt G., Mosquera C., Nápoles G., & et al., A review on the long short-term memory model. Artif Intell Rev, 53, 5929–5955, 2020.
[41] Reddy D., & Prasad P., Prediction of vegetation dynamics using NDVI time series data and LSTM. Modeling Earth Systems and Environment, 4(1), 409-419, 2018.
[42] Moghar A., & Hamiche M., Stock market prediction using LSTM recurrent neural network. Procedia Computer Science, 170, 1168-1173, 2020.
[43] Rajagukguk R., Ramadhan R., & Lee H. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power, Energies, 13(24), 6623, 2020.
[44] Zeydalinejad N., Artificial neural networks vis-à-vis MODFLOW in the simulation of groundwater: A review, Modeling Earth Systems and Environment, 1-22, 2022.
[45] Zhang B., Zou G., Qin D., Lu Y., Jin Y., & Wang H., A novel Encoder-Decoder model based on read-first LSTM for air pollutant prediction. Science of The Total Environment, 765, 144507, 2021.
[46] Park S., Kim B., Kang C., Chung C., & Choi J., Sequence-to-sequence prediction of vehicle trajectory via LSTM encoder-decoder architecture. IEEE Intelligent Vehicles Symposium, 1672-1678, 2018.
[47] Ellis M., & Chinde V., An encoder–decoder LSTM-based EMPC framework applied to a building HVAC system, Chemical Engineering Research and Design, 160, 508-520, 2020.
[48] Wang Y., Huang M., Zhu X., & Zhao L., Attention-based LSTM for aspect-level sentiment classification. In Proceedings of the 2016 conference on empirical methods in natural language processing, 606-615, 2016.
[49] Kim S., & Kang M. (2019). Financial series prediction using Attention LSTM. arXiv preprint arXiv, 1902-10877.
[50] Liu G., & Guo J., Bidirectional LSTM with attention mechanism and convolutional layer for text classification. Neurocomputing, 325-338, 2019.
[51] Biswas S., & Sinha M., Performances of deep learning models for Indian Ocean wind speed prediction, Modeling Earth Systems and Environment, 7(2), 809-831, 2021.
[52] Li Y., Zhu Z., Kong D., Han H., & Zhao Y., EA-LSTM: Evolutionary attention-based LSTM for time series prediction. Knowledge-Based Systems, 181, 104785, 2019.
[53] Xiao Y., Yin H., Zhang Y., Qi H., Zhang Y., & Liu Z., A dual-stage attention-based Conv-LSTM network for spatio-temporal correlation and multivariate time series prediction. International Journal of Intelligent Systems, 36(5), 2036-2057, 2021.

DOI: https://doi.org/10.54216/JISIoT.080202
Received: August 15, 2021 Accepted February 10, 2023