Hyper-parameter tuning for long short-term memory (LSTM) algorithm to forecast a disease spreading

T. Makarovskikh
South Ural State University
Chelyabinsk, Russia
makarovskikh.t.a@susu.ru

M. Abotaleb
South Ural State University
Chelyabinsk, Russia
abotalebmostafa@bk.ru

***Abstract*—Deep learning, artificial intelligence, and machine learning are ways for technologies to support decision-making in real-time to control the spread of the pandemic, which can help healthcare organizations. This study aims to investigate hyper-parameter tuning for Long Short-Term Memory to forecast Covid-19 infection cases in the Russian Federation by pick the best loss function, activation function, number of epochs, number of neurons in a cell, and optimizer to minimize the error in addition to a good fit for the model where the performance of the model is good on both the training and validation sets. Based on machine learning long short-term memory (LSTM), which has the advantage of analyzing relationships among time series data through its memory function, we propose a forecasting method for daily Covid-19 infection cases based on bidirectional LSTM. In the meanwhile, we use about 10 different forecasting models to forecast the daily Covid-19 infection cases one by one. Moreover, the results of these models are analyzed and compared. Firstly, we concluded that bidirectional long short-term memory to efficiently extract features from the data, which are the items of the previous day. And then, we adopt Bidirectional LSTM to predict the daily infection cases with the extracted feature data. We used BiLSTM to forecast Covid-19 spreading in the Russian Federation for one month. The prediction results were expected. According to the experimental results, the Bidirectional LSTM can provide a reliable daily Covid-19 infection case forecast in Russia with the highest prediction accuracy.**

Keywords*— *Hyper-parameter, LSTM, BILSTM. ConvLSTMs, LSTMs, Machine learning, time series analysis, forecasting.

1. INTRODUCTION

With the increasing number of Covid-19 infections and the recognition of new Omicron variant, it is necessary to track new infection cases in order to control and monitor meetings, events, and provide assistance in places where the virus is rapidly spreading. Covid-19 has the ability to be the most global epidemic of this decade [1]. For forecasting the global number of Covid-19 infection cases, [2] developed a novel hybrid forecasting model, CNN–LSTM. While the proposed model was compared to 17 baseline models, CNN–LSTM outperformed them all, with the lowest root mean square error (RMSE). There is study for forecasting infection cases in India and Chennai accurately using deep learning models. The authors concluded that stacked LSTM (LSTMs) outperform the ARIMA, LSTM, and Prophet [3]. The LSTM model is the best performer for forecasting cumulative infection cases for 7 and 30 days [4]. A study was held in Egypt, Kuwait, and Saudi Arabia to forecast Covid-19 cases, and authors concluded that LSTM had the best performance in infection cases for the three countries, and GRU had the best performance in death cases for Egypt and Kuwait [5]. For infection and recovered cases, the LSTM-CNN achieved improved performance with an average mean absolute percentage error (MAPE) among others [6]. Empirical studies conducted show that deep learning-based algorithms such as ANN and LSTM outperform traditional-based algorithms such as the ARIMA model [7]. The aim of this research is to investigate hyper-parameter tuning for long short-term memory (LSTM) models to forecast Covid-19 infection cases in Russia. After we compared the mean absolute error (MAE), RMSE, and R-squared (R^2), we found that the BLSTM model is the best model for forecasting Covid-19 infection cases in Russia for the current time, as it achieved the least error in MAE and RMSE as well as the highest R^2.

2. PROPOSED FRAMEWORK

The mechanism of the proposed framework for forecasting of time series $x_1, x_2, x_3, \ldots, x_n$ is following.

Step 1. Data preparation. We have some problems with lengthy sequences in the database [8]. The first problem is that it takes a long time for training and needs a lot of memory, and the second problem is getting a poor learned model caused by back-propagating lengthy sequences. So, it is important to prepare and preprocess the data before it is imported to neural networks. Data normalization and standardization are two techniques used in the data preparation phase. We used data standardization, which is realized as a scaling approach to set the mean and standard deviation to 0 and 1, respectively [9].

Step 2. Split dataset into training, validation, and testing sets. Daily Covid-19 infection cases in the Russian Federation Data from the World Health Organization were collected from the January, 3, 2020 to December, 31, 2021. We test our model using 5% of this dataset (37 days). The dataset is further prepared such that the first 95% of the datasets are used for training and the last 5% of the datasets are used for testing purposes. The training set is used to train and improve the models, and we divided the 20% of training data for validation purposes to diagnose overfitting and underfitting, whereas the test set is used to assess the performance of the model. The algorithm and the daily covid 19 infection case data can be obtained from [10].

Step 3. Models' development. In this step, we run our algorithm for LSTM, LSTMs (stacked LSTM), bidirectional LSTM, LSTMs (stacked BILSTM), convolutional neuron networks (Conv), ConvLSTMs and other forecasting models.

Step 4. Evaluate the forecasting using data from [11]. In order to evaluate the forecasting effect of hyper-parameter tuning for machine learning models, the mean absolute error (MAE), root mean square error (RMSE), and R-square (R2) are used as the evaluation criteria for the methods (see Table 1).

3. Results

Таблица I. Table 1. Comparison of 10 methods evaluation testing Covid 19 daily infection cases in Russia

Method	MAE	RMSE	R^2
LSTM	621.432	753.053	0.962
Stacked LSTM	1025.523	1273.222	0.893
LSTM 2 window	711.051	857.457	0.949
LSTM 3 window	642.859	776.476	0.957
LSTM 4 window	604.986	741.646	0.960
LSTM 10 window	640.832	742.132	0.949
BILSTM	**552.889**	**658.607**	**0.971**
Stacked BILSTM	965.569	1192.145	0.906
CONV	849.763	921.963	0.578
ConvLSTMs	1189.480	1350.587	0.094

The bold values present the lowest error values of (MAE, RMSE, and R^2).

These values show that the proposed approach outperforms the baseline models based on the testing data. After using the processed training set data to train LSTM, Stacked LSTM (LSTMs), LSTM 2 window, LSTM 3 window, LSTM 4 window, LSTM 10 window, BILSTM, Stacked BILSTM, CONV, and ConvLSTMs, respectively, the model completed by training is used to predict the test set data, and the real value is compared with the predicted. From Table 1 the MAE and RMSE of ConvLSTMs are the largest and R2 is the smallest, while the MAE and RMSE of bidirectional LSTM (BI-LSTM) are the smallest, R^2 is the largest, and the closest is 1.

The results show the performance of the BI-LSTM developed model is the best among the other 10 methods. In terms of forecasting accuracy, MAE and RMSE are the smallest among the considered models. Hence, this method has high forecasting accuracy for the investigated dataset. In terms of forecasting performance, the R^2 of BI-LSTM can be compared with the other methods. Therefore, the BI-LSTM proposed in this paper is superior to the other considered comparative models in terms of fitting degree and error value.

In our research, the role of BiLSTM is crucial, where prediction depends not only on the previous input but also the future input. Because of the nature of daily Covid-19 infection cases in Russia, BiLSTM outperforms the rest of the models. The network needs to increase the amount of information available to the network and does so through two LSTMs that work in two different directions, one taking the input in a forward direction and the other in a backwards direction, so BiLSTM works effectively. We designed the BiLSTM to look at the previous day to forecast the next day with the minimum error.

4. Conclusion

Despite that the spread of Covid-19 is recognized as an impossible process for forecasting, the available data sets can be used to study the properties of various predictive models, in particular, for the development of artificial intelligence methods. These methods can be used to forecast not only epidemic data. Developed approaches to improve the accuracy of forecasting can increase the quality of simulation for other processes, such as socio-economic, technological, physical, etc. In our case, the using more than one time step in BiLSTM lead to increasing of error. Hence, we used only one time step, meaning that the forecast for infection cases for the next day depends on the previous day. We used BiLSTM to forecast Covid-19 spreading in the Russian Federation for one month [11]. The prediction results were expected. In the absence of a new Covid-19 variant, the current wave, which began in the first days of January, will go down in about a month, which is currently observed. If a new variant appears, another wave is possible most probably before the end of summer. Since the situation with virus spreading changes daily, we recommend rerunning our proposed framework every day to get results with lower error. This forecasting method not only provides a new research idea for infection case forecasting but also provides practical experience for scholars. On the other hand, it provides a reference for the health sector in Russia, in particular, and the world health organization WHO, and in general for the health sectors in other countries. In future research, We will develop architecture of the Attention LSTM model and Attention BiLSTM model to minimize the errors and compare the errors with those of this research.

Acknowledgements

The work was supported by the Ministry of Science and Higher Education of the Russian Federation (government order FENU-2020-0022).

References

[1] Kannan, S.P. COVID-19 (Novel Coronavirus 2019)-recent trends / S.P. Kannan, P.S. Ali, A. Sheeza, K. Hemalatha // Eur Rev Med Pharmacol Sci. – 2020. – Vol. 24(4). – P. 2006-2011.

[2] Zuhaira, Z.M. COVID-19 pandemic forecasting using CNN-LSTM: a hybrid approach / Z.M. Zuhaira, N.M. Alturki // Journal of Control Science and Engineering. – 2021. – Vol. 2021.

[3] Devaraj, J. Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant? / J. Devaraj, R.M. Elavarasan, R. Pugazhendhi, G.M. Shafiullah, S. Ganesan, A.K. Jeysree, I.A. Khan, E. Hossain // Results in Physics. – 2021. – Vol. 21. – P. 103817.

[4] Marzouk, M. Deep learning model for forecasting COVID-19 outbreak in Egypt / M. Marzouk, N. Elshaboury, A. Abdel-Latif, S. Azab // Process Safety and Environmental Protection. – 2021. – Vol. 153. – P. 363-375.

[5] Omran, N.F. Applying deep learning methods on time-series data for forecasting COVID-19 in Egypt, Kuwait, and Saudi Arabia / N.F. Omran, S.F. Abd-el Ghany, H. Saleh, A. Abdelmgeid, A. Gumaei, M. Al-Rakhami // Complexity. – 2021.- Vol. 2021.

[6] Dairi, A. Comparative study of machine learning methods for COVID-19 transmission forecasting // A. Dairi, F. Harrou, A. Zeroual, M.M.Hittawe, Y. Sun // Journal of Biomedical Informatics. – 2021. – Vol. 118. – P. 103791.

[7] Abdollahi, J. Modeling and forecasting Spread of COVID-19 epidemic in Iran until Sep 22, 2021, based on deep learning / J. Abdollahi, A.J. Irani, B. Nouri-Moghaddam // ArXiv preprint: 2103.08178, 2021.

[8] Khan, S.D.Toward smart lockdown: a novel approach for COVID-19 hotspots prediction using a deep hybrid neural network / S.D. Khan, L. Alarabi, S. Basalamah // Computers. – 2020. – Vol. 9(4). – P. 99.

[9] Ali, P.J.M. Data normalization and standardization: a technical report / P.J.M. Ali, R.H. Faraj, E. Koya // Mach Learn Tech Rep. – 2014. – Vol. 1(1). – P. 1-6.

[10] Abotaleb, M. Hyper-parameter Tuning for Long Short-Term Memory (LSTM) Algorithm to Forecast a Disease Spreading / M. Abotaleb, T. Makarovskikh. URL: https://github.com/ abotalebmostafa11/Hyper-parameter-Tuning.

[11] Coronavirus: statistics. URL: https://yandex.ru/covid19/stat.