

UDC 519.6

DOI: 10.22363/2658-4670-2021-29-1-22-35

Study of the impact of the COVID-19 pandemic on international air transportation

Eugeny Yu. Shchetinin

*Financial University under the Government of Russian Federation
49, Leningradsky Prospect, Moscow, 125993, Russian Federation*

(received: February 20, 2021; accepted: March 12, 2021)

Time Series Forecasting has always been a very important area of research in many domains because many different types of data are stored as time series. Given the growing availability of data and computing power in the recent years, Deep Learning has become a fundamental part of the new generation of Time Series Forecasting models, obtaining excellent results.

As different time series problems are studied in many different fields, a large number of new architectures have been developed in recent years. This has also been simplified by the growing availability of open source frameworks, which make the development of new custom network components easier and faster.

In this paper three different Deep Learning Architecture for Time Series Forecasting are presented: Recurrent Neural Networks (RNNs), that are the most classical and used architecture for Time Series Forecasting problems; Long Short-Term Memory (LSTM), that are an evolution of RNNs developed in order to overcome the vanishing gradient problem; Gated Recurrent Unit (GRU), that are another evolution of RNNs, similar to LSTM.

The article is devoted to modeling and forecasting the cost of international air transportation in a pandemic using deep learning methods. The author builds time series models of the American Airlines (AAL) stock prices for a selected period using LSTM, GRU, RNN recurrent neural networks models and compare the accuracy forecast results.

Key words and phrases: neural networks, financial forecasting, deep learning, international air travel

1. Introduction

In 2020, there was a significant drop in quotations of American Airlines (AAL) associated with the COVID-19 pandemic and a record-breaking decrease in the number of air travel in the world. The generally accepted econometric methods of modeling and forecasting financial time series in these conditions turned out to be ineffective for making even short-term forecasts [1], [2]. In the present paper, methods for modeling and forecasting

© Shchetinin E. Yu., 2021



This work is licensed under a Creative Commons Attribution 4.0 International License

<http://creativecommons.org/licenses/by/4.0/>

international air traffic in the 2019–2020 pandemic are explored using recurrent neural networks with different architectures. As an object of research, the day quotes of the American company AAL, traded on the NASDAQ exchange, were selected; data from September 27, 2005 to September 30, 2020 from the information portal Yahoo Finance [3] were taken. The shares of this US company were selected due to its leading positions in the international air transportation market, high values of the trading turnover on the NASDAQ exchange, which in turn provides liquidity and shows investor interest in this exchange commodity [4]. Using the example of the value of AAL shares, we will try to build a reliable forecast using deep learning methods, in particular, recurrent neural networks [5]–[7].

2. Pre-processing of input data

As input data for the neural network model, we will take a sequence consisting of the following values:

- Opent-1 — opening price for the previous period;
- Lowt-1 — the minimum price for the previous trading day;
- Hightt-1 — the maximum price for the previous trading day;
- Volumet-1 — the amount of shares sold and bought for the previous trading day;
- Closet-1 — closing price for the previous trading day.

Based on the input data, neural networks will generate an output value that can be interpreted as the predicted value of the closing quotation today. For the correct operation of neural networks, it is necessary to normalize the data within the limits of $[0 : 1]$, as well as create training and test samples in the ratio 80:20 from the initial data having the dimension 3636. Thus, 2909 observations for the training sample and 727 observations for the test sample were obtained. The table 1 shows a fragment of the input data.

It is necessary to remove the Date and Adj Close columns from the received data. The table 2 presents descriptive statistics of input data. It is seen that the average closing price is \$27.13 and the standard deviation is \$16.74.

To study the statistical properties of the data further, let us build scatter diagrams of the profitability of the opening price and the closing price, as well as the profitability of the closing price shifted by one lag, and the closing price today. To calculate the profitability, we will use the following formula [8]–[10]:

$$R = \frac{y_t}{y_{t-1}} - 1, \quad (1)$$

where R is the profitability; y_{t-1} is the previous observation value; y_t are the values for the current time period.

The scatter diagram of the profitability of the opening and closing prices is shown in the figure 1.

The figure 1 shows that there is no correlation between the variables under consideration. Next, we will construct a histogram of the distribution of the profitability of closing prices (figure 2). Obviously, most of the observations are in the range from -0.1 to 0.1 . This means that in most observations, the price changed from -10% to 10% in one period.

Table 1

A fragment of the input data

Observation number	Date	Open, \$	High, \$	Low, \$	Close, \$	Adj Close ¹ , \$	Volume, \$
0	27.09.2005	21.05	21.4	19.1	19.3	18.19	961200
1	28.09.2005	19.3	20.53	19.2	20.5	19.33	5747900
2	29.09.2005	20.4	20.58	20.1	20.21	19.05	1078200
3	30.09.2005	20.26	21.05	20.18	21.01	19.81	3123300
4	3.10.2005	20.9	21.75	20.9	21.5	20.27	1057900
...
3634	6.03.2020	15.02	17.12	14.8	15.97	15.97	54505000
3635	9.03.2020	14.87	15.79	14.46	14.75	14.75	42558000
3636	10.03.2020	15.82	17.67	14.61	17	17	56858200

Table 2

Descriptive statistics of input data

Indicators	Open	High	Low	Close	Volume
Total number of observations	3637	3637	3637	3637	3.637e+03
Mean value, \$	27.15	27.64	26.64	27.13	7.603118e+06
Standard deviation, \$	16.74	16.95	16.53	16.74	6.070650e+06
Minimal value, \$	1.81	2.03	1.45	1.76	1.385e+05
25% percentile, \$	9.57	9.81	9.32	9.58	4.1782e+06
50% percentile, \$	29.9	30.48	29.28	29.89	6.5025e+06
75% percentile, \$	41.74	42.24	41.02	41.68	9.5455e+06
Maximum value, \$	62.7	63.27	62	62.95	1.377672e+08

To test the hypothesis about whether the distribution of the closing price profitability is a special case of the normal distribution, we use the Shapiro–Wilk and Jarque–Bera tests. The Jarque–Bera test rejected the null hypothesis at a significance level of $\alpha = 0.05$. The results of the Shapiro–Wilk test and the Jarque–Bera test coincided. This means that the profitability of closing prices has a distribution that is different from the normal one.

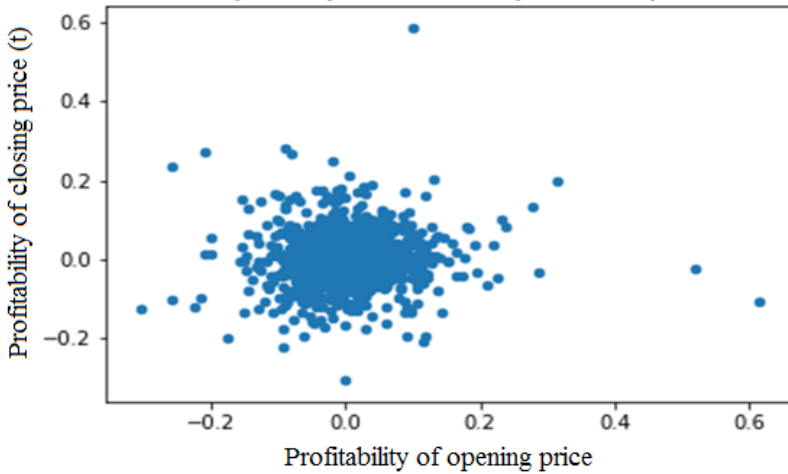


Figure 1. Scatter diagram of opening and closing prices

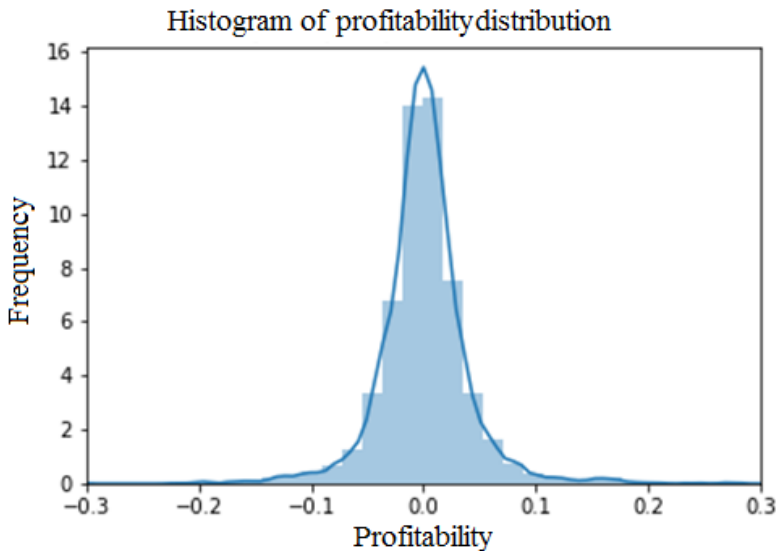


Figure 2. Distribution of closing price profitability

To check the stationarity of the profitability series, we will use the Dickey–Fuller test, which is one of the unit root tests. A time series has a unit root if its first differences form a stationary series, i.e. a series whose properties do not change over time. This condition is written as $y_t \sim I(1)$ if the series of the first differences $\Delta y_t = y_t - y_{t-1}$ is a stationary series $\Delta y \sim I(0)$ [11]. If the time series has a unit root, then it is not a stationary time series, but an integrated first-order time series [12]–[14]. As one would expect, the observed time series has no unit roots and, therefore, is stationary. For the convenience of using the input data, we will normalize them. The results are presented in the table 3.

Table 3

Normalized raw data

Number of observation	Open	High	Low	Close	Volume
0	0.31598	0.316297	0.291495	0.286648	0.005978
1	0.287239	0.30209	0.293146	0.306259	0.040757
2	0.305305	0.302907	0.30801	0.30152	0.006828
3	0.303005	0.310581	0.309331	0.314594	0.021687
4	0.313516	0.322012	0.321222	0.322602	0.00668
...
3634	0.216949	0.246408	0.220479	0.232227	0.395023
3635	0.214485	0.22469	0.214864	0.21229	0.308217
3636	0.230087	0.255389	0.217341	0.24906	0.412121

Next, we turn to the description of the main models of recurrent neural networks and their application in the analysis of financial time series.

3. Basic models of deep neural networks for simulation of financial time series

3.1. Basic recurrent neural network

The architecture of the proposed basic recurrent neural net (RNN) is as follows. A matrix with a dimension of 1 by 5 is fed to the input of the neural network, then the values are transferred to a recurrent layer with 25 neurons, after which the operation is repeated and the values are again fed to the recurrent layer with 25 neurons. At the penultimate step, the values are transferred to an aggregating layer with a dimension of 5 neurons, the result is displayed as a predicted value. Hidden layers have a hyperbolic tangent as an activation function. This activation function is nonlinear, which allows layers to be linked, i.e. combines them, because the combination of non-linear functions is also a non-linear function. Another advantage of the hyperbolic tangent function is that it is a smooth function, and this function is not binary and takes values in the range $(-1, 1)$, which eliminates overloading from large values. The hyperbolic tangent is very similar to the sigmoid with the difference that it has a larger gradient than the sigmoid. On the aggregate layer, a linear function is used as the activation function. The proposed neural network model, all procedures for its training and testing were implemented in the Keras library of the Python programming language [15].

The mean squared error (MSE) will be used as the loss function, and the optimization is performed using the Adam algorithm. The epoch parameter of the fit function reflects how many times the sample is passed through the neural network, in this case epoch = 150. The batch_size parameter is responsible for the size of the so-called batch. In cases where the training sample is too large, there is a need to divide it into parts. These parts are called batches. Thus, the training set with 2109 observations is divided into 210 batches with a size of 10, except for the last one with 9 observations. Thus, 210 iterations were required to pass one epoch.

Due to the tendency of recurrent neural networks to overfit, it is necessary to apply various regularization algorithms [10], [16]. As such an algorithm, the early stop method is used, which tracks the amount of losses. If during 20 epochs the improvement is less than 0.000002, then the training of the model will be stopped. The graph of the loss function on the training sample is shown in the figure 3.

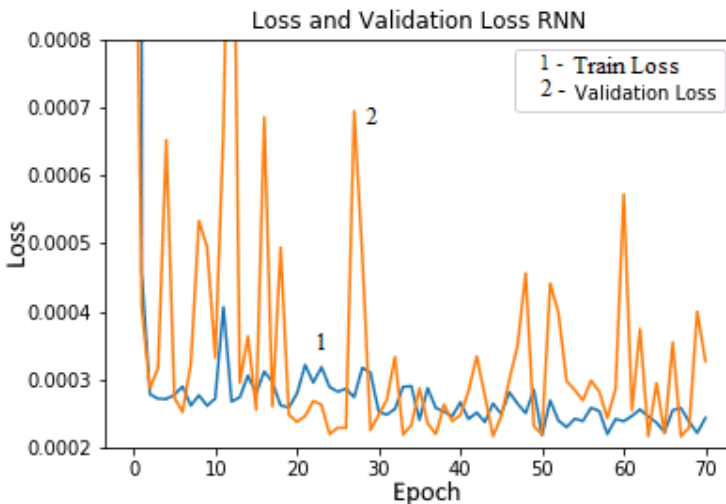


Figure 3. Plot of the RNN learning loss function:
1 — train loss; 2 — validation loss

After checking and training the neural network, we will construct a forecast of closing prices for the test sample. For a better visual appearance, the predicted values are shifted ten units up. Let us display the forecast of the last 50 observations of the test sample for a more accurate visual examination (figure 4). It can be seen from the figure that the neural network predicts closing prices closely enough.

3.2. Neural network with a gated recurrent unit

A recurrent neural network based on a cell architecture with a gated recurrent unit (GRU) repeats the structure of the RNN model of a recurrent network. The input layer takes the values of a matrix with a dimension of 1 by 5. Then, recurrent layers with 25 neurons and a hyperbolic tangent as an activation function are sequentially accepted and processed.

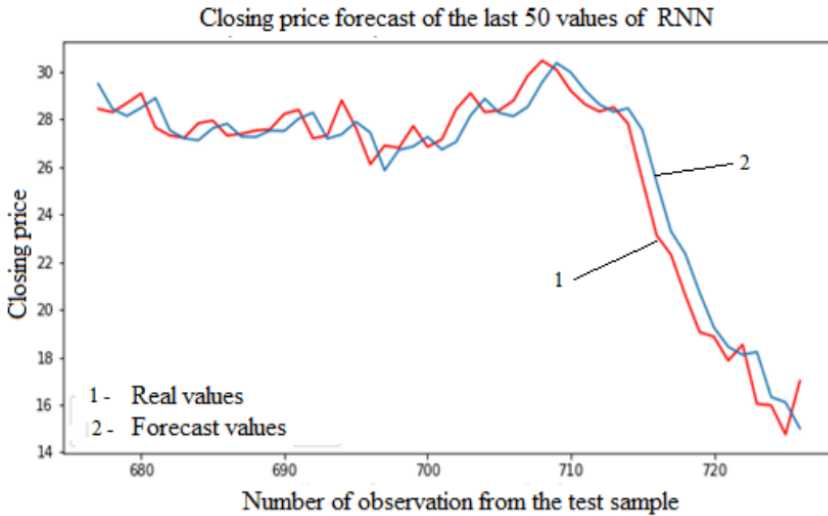


Figure 4. Forecast of the closing price for the last 50 values of RNN model network:
1 — real values stock price; 2 — forecast price

The aggregating layer has 5 neurons with a linear activation function. After processing by the last layer, the predicted value is supplied. It should be noted that the default activation function for layers with the GRU architecture is the hyperbolic tangent [16], [17]. The loss plot for the GRU recurrent neural network is shown in the figure 5.

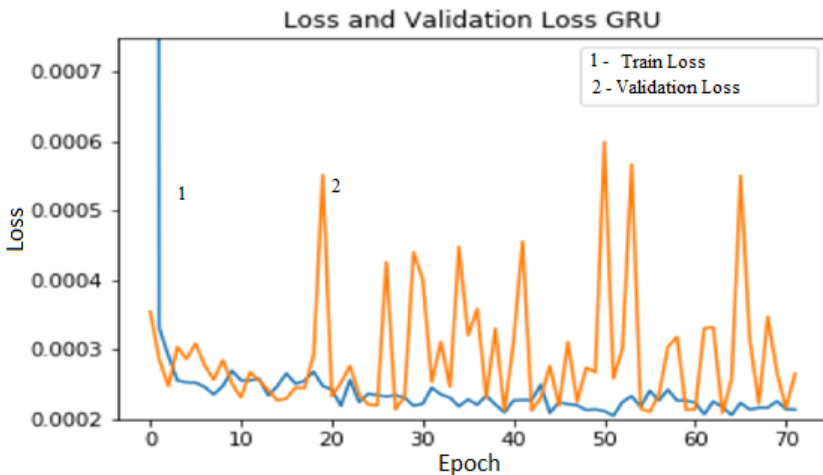


Figure 5. Loss plot for GRU model network:
1 — train loss function; 2 — validation loss function

The early stop regularization terminated the training of the neural network to prevent overfitting at epoch 72. The plot of predicted closing prices for all observations of the model of a recurrent neural network with the GRU architecture is shown in the figure 6. As in the case of the RNN, to facilitate

visualization the predicted values have been shifted ten units upward. It is also worth noting that the neural network accurately reproduced the closing price behavior. For a detailed consideration, we take the last 50 values of the test sample and display them in the figure 7.

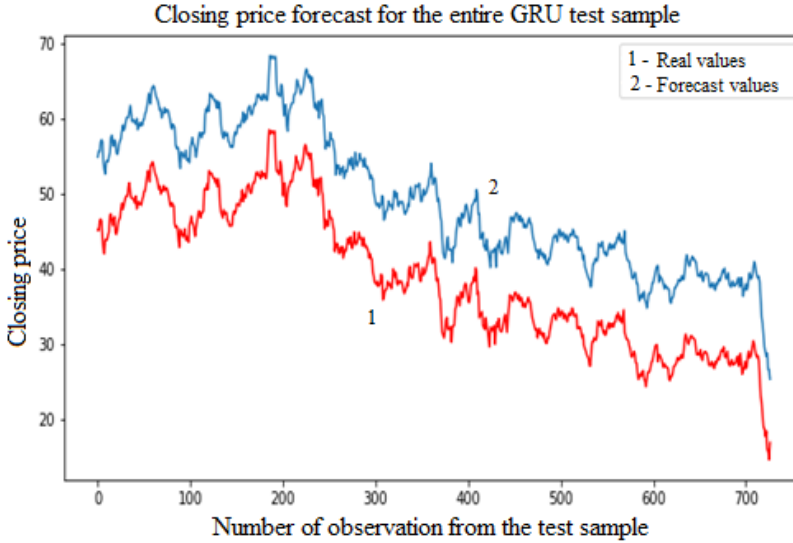


Figure 6. Closing price forecast for the entire GRU test sample:
1 — real values stock price; 2 — forecast stock price

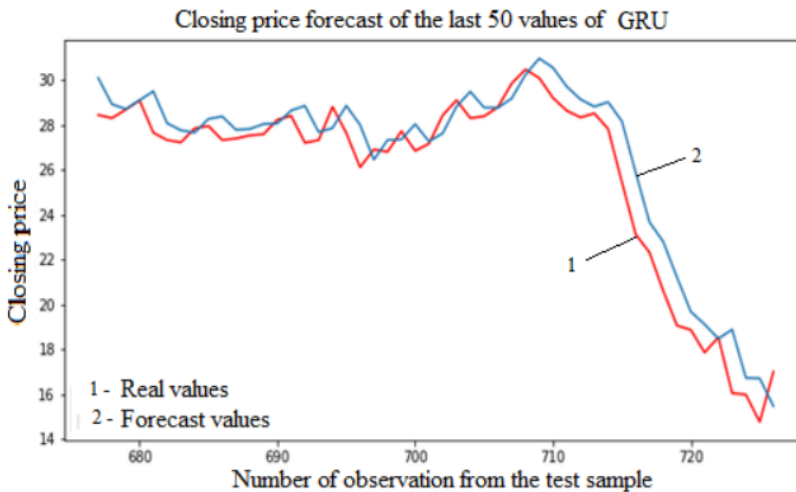


Figure 7. Closing price forecast for the last 50 values of GRU:
1 — real values of stock prices; 2 — forecast stock prices

The mean square forecast error and the R^2 index have the following values: $MSE = 0.9953$, $R^2 = 0.9885$.

3.3. Neuron network with long short-term memory (LSTM)

Just like the previous networks, constructively recurrent neural network with long short-term memory (LSTM) will repeat the previous values. An input that accepts a 1-by-5 matrix transmits information to two recurrent layers with 25 neurons per layer and a hyperbolic tangent as an activation function. Then an aggregating layer of five neurons with a linear activation function passes the value to the output layer.

The closing price prediction plot calculated using a recurrent neural network with the LSTM architecture is shown in the figure 8.

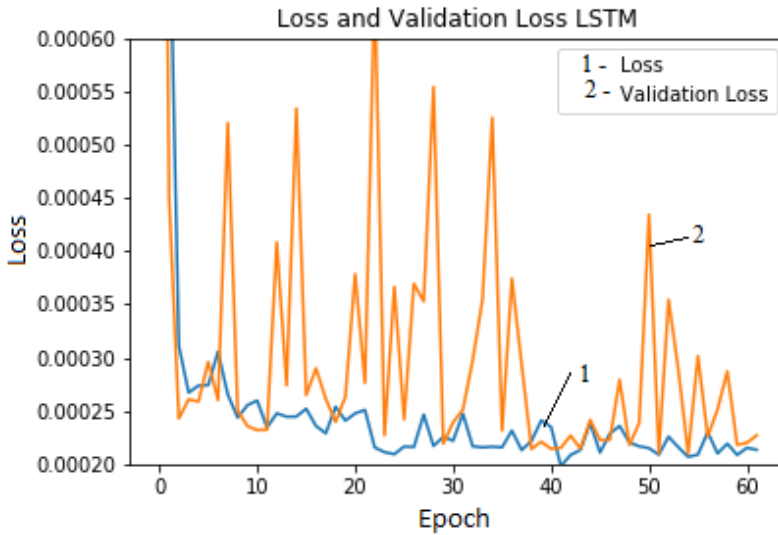


Figure 8. LSTM model network loss plot:
1 — train loss function; 2 — validation loss function

Forecasted values are shifted ten points. Based on the plot, we can conclude that the neural network under consideration predicts the required values quite accurately.

The forecast of the closing price for the entire LSTM test sample and for the last 50 values is shown in figures 9 and 10 respectively. For this recurrent neural network, $MSE = 0.8508$, $R^2 = 0.99$.

Let us display a comparative plot of losses during training of various constructions and architectures of the considered neural networks (figure 11). Note that the RNN recurrent neural network demonstrated the highest loss rates on the training set. Except for separately taken random epochs, its loss value was greater than that of the rest. LSTM and GRU recurrent neural networks have close values of losses on the training set. It is worth noting that the early stopping algorithm worked for all types of recurrent neural networks. For the RNN model, the algorithm stopped training at 71 epochs, for GRU — at 72. The least number of epochs – 62 — was required to train the neural network built using the LSTM architecture.

The table 4 shows the values of the mean square error and the coefficient R^2 for all constructed neural networks.

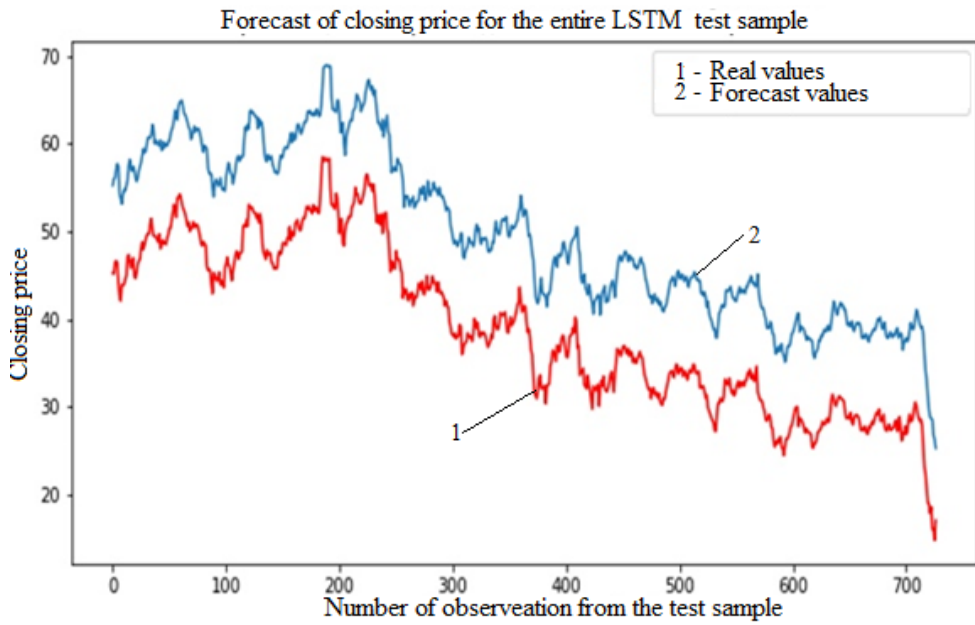


Figure 9. Forecast of the closing price for the entire LSTM model network test sample:
 1 — Real values stock prices; 2 — forecast stock prices

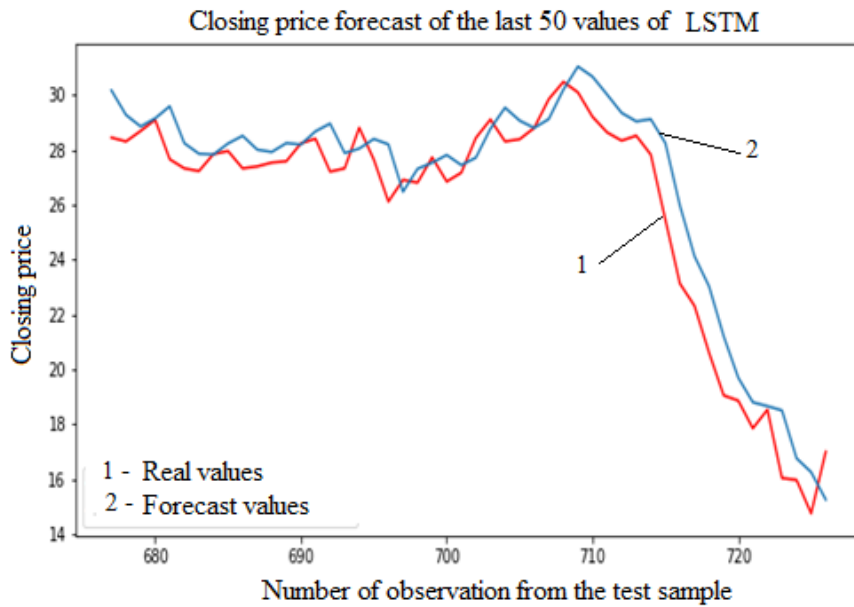


Figure 10. Forecast of the closing price of the last 50 values of LSTM model network:
 1 — Real values stock prices; 2 — forecast stock prices

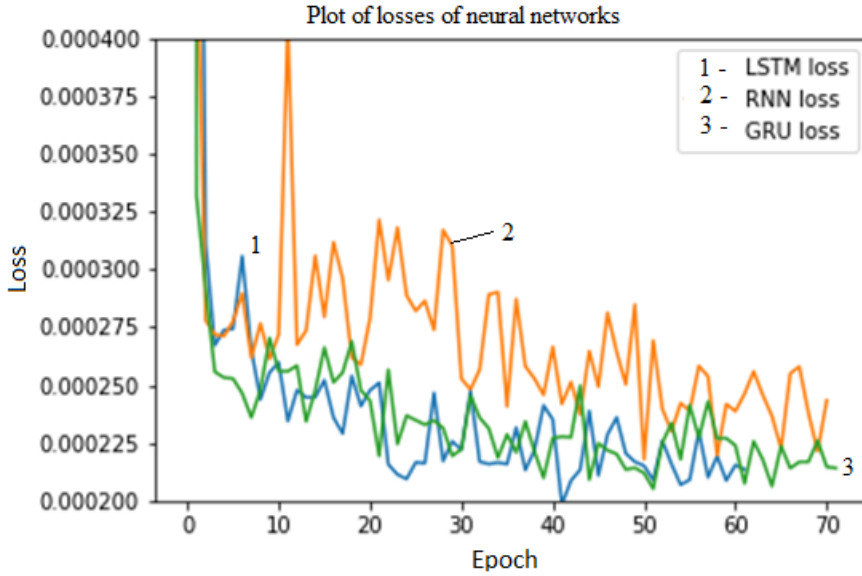


Figure 11. Plot of losses for different models of neural networks:
 1 — LSTM model loss function; 2 — RNN model loss function; 3 — GRU model loss function

Table 4

Values of MSE и R^2 for all constructed neural networks

Neural network	MSE	R^2
RNN	1.2232	0.9858
GRU	0.9953	0.9885
LSTM	0.8508	0.9901

4. Discussion of results of computer experiments

In the process of investigating the impact of the COVID-19 pandemic on AAL stock quotes, recurrent neural network models were built with various architectures, such as cells with long short-term memory LSTMs, cells with gated recurrent unit GRU, and a basic recurrent network. The analysis of the constructed models was carried out, as well as the comparison of the results on the training and test data. During the analysis, it was found that the neural network with long short-term memory cells (LSTM) coped best with the task of predicting the data under study.

Summing up, we can say that all networks have shown a satisfactory result, but they predict the price with a certain delay, which may entail unplanned financial losses. In view of this, it can be concluded that these models are not suitable for carrying out short-term operations in the financial market,

are not able to serve as an indicator that helps to improve the efficiency of a trading strategy and cannot be used for risk management tasks.

5. Conclusion

The purpose of the article was to investigate the quality of various neural network models that predict the closing price of a stock. In the course of the study, sufficiently accurate results of modeling and forecasting financial time series for the intraday closing prices of shares of the American airline ALL were obtained, which confirmed the effectiveness of using the proposed models of deep neural networks. However, in the context of the practical application of the developed models, it is necessary to take into account time delays in obtaining forecast results, as well as the horizon of financial forecasting.

References

- [1] J. D. A. Hamilton, *The time series analysis*. Princeton New Jersey: University Press, 1994.
- [2] C. Brooks, *Introductory econometrics for finance*. Cambridge: Cambridge University Press, 2019.
- [3] “American Airlines Group Inc. (AAL),” URL: <https://finance.yahoo.com/quote/AAL/>. Availabel: 2020-11-25, 2020.
- [4] E. Y. Shchetinin, “On a structural approach to managing a company with high volatility of indicators [K analizu effektivnosti biznesa v usloviyah vysokoj izmenchivosti ego finansovyh aktivov],” *Finansy i kredit*, vol. 14, no. 218, pp. 39–41, 2006, [in Russian].
- [5] J. Vander Plas, *Python Data Science Handbook*. Sebastopol, CA: O’Reilly Media, 2016.
- [6] W. Richert and L. P. Coelho, *Building Machine Learning Systems with Python*. Birmingham: Packt, 2013.
- [7] C. Bishop, *Pattern recognition and machine learning*. Berlin, Germany: Springer-Verlag, 2006.
- [8] E. Y. Shchetinin, “Modeling the energy consumption of smart buildings using artificial intelligence,” in *CEUR Workshop Proceedings*, vol. 2407, 2019, pp. 130–140.
- [9] M. Mudelsee, “Trend analysis of climate time series: A review of methods,” *Earth-Science Reviews*, vol. 190, pp. 310–322, 2019. DOI: 10.1016/j.earscirev.2018.12.005.
- [10] C. Chen, J. Twycross, and J. M. Garibaldi, “A new accuracy measure based on bounded relative error for time series forecasting,” *PLOS ONE*, vol. 12, no. 3, pp. 1–23, Mar. 2017. DOI: 10.1371/journal.pone.0174202.
- [11] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. Melbourne, Australia: OTexts, 2018.

- [12] A. Ghaderi, B. M. Sanandaji, and F. Ghaderi. “Deep forecast: deep learning-based spatio-temporal forecasting.” arXiv: 1707.08110 [cs.LG]. (2017).
- [13] S. B. Taieb, A. Sorjamaa, and G. Bontempi, “Multiple-output modeling for multi-step-ahead time series forecasting,” *Neurocomput.*, vol. 73, no. 10, pp. 1950–1957, 2010. DOI: 10.1016/j.neucom.2009.11.030.
- [14] R. Sen, H.-F. Yu, and I. S. Dhillon, “Think globally, act locally: a deep neural network approach to high-dimensional time series forecasting,” in *Advances in neural information processing systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett, Eds., vol. 32, Curran Associates, Inc., 2019.
- [15] “Keras,” URL: <https://www.keras.io>. Availabel: 2020-11-25, 2020.
- [16] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. Cambridge: The MIT Press, 2016.
- [17] S. Galeshchuk and S. Mukherjee, “Deep networks for predicting direction of change in foreign exchange rates,” *Intelligent Systems in Accounting, Finance and Management*, vol. 24, no. 4, pp. 100–110, 2017. DOI: 10.1002/isaf.1404.

For citation:

E. Yu. Shchetinin, Study of the impact of the COVID-19 pandemic on international air transportation, *Discrete and Continuous Models and Applied Computational Science* 29 (1) (2021) 22–35. DOI: 10.22363/2658-4670-2021-29-1-22-35.

Information about the authors:

Shchetinin, Eugeny Yu. — Doctor of Physical and Mathematical Sciences, Lecturer of Department of Mathematics (e-mail: riviera-molto@mail.ru, ORCID: <https://orcid.org/0000-0003-3651-7629>)

УДК 519.6

DOI: 10.22363/2658-4670-2021-29-1-22-35

Исследование влияния пандемии COVID-19 на международные авиаперевозки

Е. Ю. Щетинин

*Финансовый университет при Правительстве Российской Федерации
Ленинградский проспект, д. 49, Москва, 125993, Россия*

Прогнозирование временных рядов играет важную роль во многих областях исследований. Вследствие растущей доступности данных и вычислительных мощностей в последние годы глубокое обучение стало фундаментальной частью нового поколения моделей прогнозирования временных рядов, получающих отличные результаты.

В данной работе представлены три различные архитектуры глубокого обучения для прогнозирования временных рядов: рекуррентные нейронные сети (RNN), которые являются наиболее известной и используемой архитектурой для задач прогнозирования временных рядов; долгая краткосрочная память (LSTM), которая представляет собой обобщённую и развитую РНС, разработанную для преодоления проблемы исчезающего градиента; закрытый рекуррентный блок (GRU), который является ещё одной эволюционной моделью РНС.

Статья посвящена моделированию и прогнозированию стоимости международных авиаперевозок в условиях пандемии с использованием методов глубокого обучения и моделей рекуррентных сетей. В работе построены модели временных рядов цен акций American Airlines (AAL) с использованием моделей рекуррентных нейронных сетей LSTM, GRU, RNN и проведён сравнительный анализ результатов точности прогноза на выбранный период. Его результаты показали эффективность применения алгоритмов глубокого обучения для оценивания точности прогнозирования временных рядов.

Ключевые слова: нейронные сети, финансовое прогнозирование, глубокое обучение, международные авиаперевозки