UNIWERSYTET TECHNOLOGICZNO-PRZYRODNICZY
IM. JANA I JĘDRZEJA ŚNIADECKICH W BYDGOSZCZY
ZESZYTY NAUKOWE NR 271
TELEKOMUNIKACJA I ELEKTRONIKA 24 (2020) 13–30

FORECASTING FUTURE VALUES OF TIME SERIES USING THE LSTM NETWORK ON THE EXAMPLE OF CURRENCIES AND WIG20 COMPANIES

Bartosz Mróz, Filip Nowicki

Faculty of Telecommunications, Computer Science and Electrical Engineering,
UTP University of Science and Technology,
Al. prof. S. Kaliskiego 7, 85-796 Bydgoszcz, Poland
e-mail: {barmro005, filnow003}@utp.edu.pl

Summary: The article presents a comparison of the RNN, GRU and LSTM networks in predicting future values of time series on the example of currencies and listed companies. The stages of creating an application which is a implementation of the analyzed issue were also shown – the selection of networks, technologies, selection of optimal network parameters. Additionally, two conducted experiments were discussed. The first was to predict the next values of WIG20 companies, exchange rates and cryptocurrencies. The second was based on investments in cryptocurrencies guided solely by the predictions of artificial intelligence. This was to check whether the investments guided by the predictions of such a program have a chance of effective earnings. The discussion of the results of the experiment includes an analysis of various interesting phenomena that occurred during its duration and a comprehensive presentation of the relatively high efficiency of the proposed solution, along with all kinds of graphs and comparisons with real data. The difficulties that occurred during the experiments, such as coronavirus or socio-economic events, such as riots in the USA, were also analyzed. Finally, elements were proposed that should be improved or included in future versions of the solution – taking into account world events, market anomalies and the use of supervised learning.

Keywords: recurrent neural network (RNN), gated recurrent unit (GRU), long short-term memory (LSTM)

1. INTRODUCTION

Both the stock exchange and the national currency are one of the most important elements of state activity. They are responsible for the stability of the economy or attracting investors. It is thanks to successful investments that man can get a lot of money. The right investments used to be associated with long hours of chart analysis, data gathering, and drawing conclusions. Today, when the world is booming in the sciences related to artificial intelligence and machine learning, this arduous process can be supported or entirely performed by computers. With the help of artificial intelligence, and more precisely LSTM networks, the properties of which are perfect for predicting time series. We decided to conduct an experiment that will show that artificial intelligence can predict the prices of three financial instruments – currencies, shares of WIG20 companies and a relatively new being on the market, i.e. cryptocurrencies – with an effectiveness higher than 90 percent. In addition, we made a program that could ultimately support investors. We also checked whether the investments guided by the predictions of such a program have a chance of high earnings.

2. METHODS

Access to many sources, such as open-source projects available on the GitHub platform, articles on the Medium portal or scientific publications, allows to properly select the best methodologies and their implementation. Taking into account the available information, properly documented libraries, the purpose of the project, as well as its assumptions – three types of neural networks can be distinguished, which were taken into account:

- Recurrent neural network (RNN),
- Gated recurrent unit (GRU),
- Long short-term memory (LSTM).

Undoubtedly, the simplest of the above-mentioned ones, taking into account both theoretical and implementation aspects, is RNN (Fig. 1 and Fig. 2). The construction itself, however, does not allow it to be used in our project. To explain why we will use the diagrams:

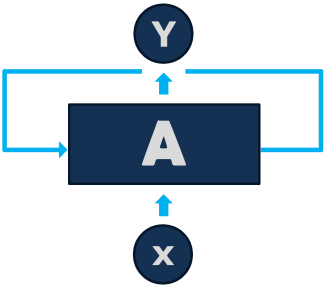

Fig. 1. Simple RNN scheme (Source: own study)

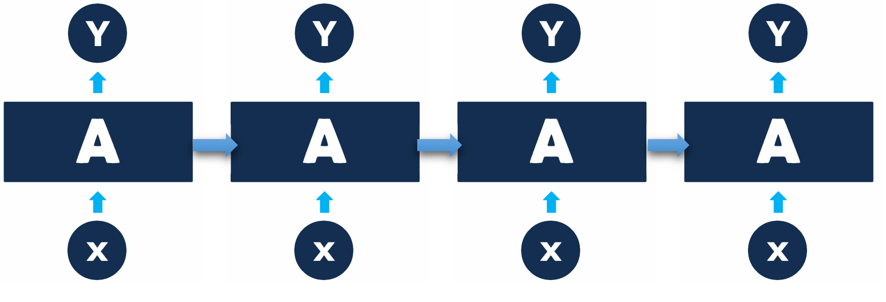

Fig. 2. A simplified diagram of the RNN network (Source: own study)

The RNN therefore has two inputs – the X parameter and the result of the previous iteration. This construction leads to a common problem called the Vanishing gradient problem. It consists in the fact that the context of the input data in the learning process gradually fades away with each subsequent iteration. In other words, we lose some important data during learning process. Why is this context so important? When building a neural network focused on features such as WIG20 stock price prediction, the most undesirable factor is data loss, which cannot happen on historical data analysis. This fact implies the obligatory maintenance of the context.

2.1. Comparison of LSTM and GRU

After the rejection of the RNN, the choice remained between LSTM and GRU. As already mentioned, we considered LSTM as the more suitable type of network for this type of project. The following schemes (Fig. 3) will be useful to justify our choice:

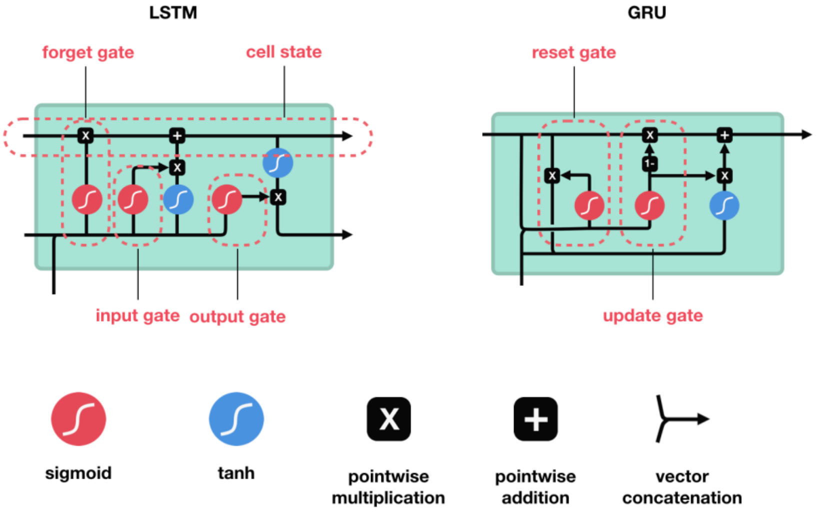

Fig. 3. LSTM (left) and GRU (right) schemas with included gateways [13]

Basing on these data, it can be concluded that the GRU has the speed advantage of operation over the LSTM – it has fewer gates, therefore the signal purely theoretically, has a shorter path to pass.

Another important issue when making a selection is the loss graph of both discussed networks. More precisely, the length of keep the context in time. Fig. 4 shows the difference between the two types of networks, supported by comparative studies conducted in the dimension of one thousand epochs by Eric Muccino. It should be noted that when choosing the right network, the epoch's intervals are an important element (Fig. 4).

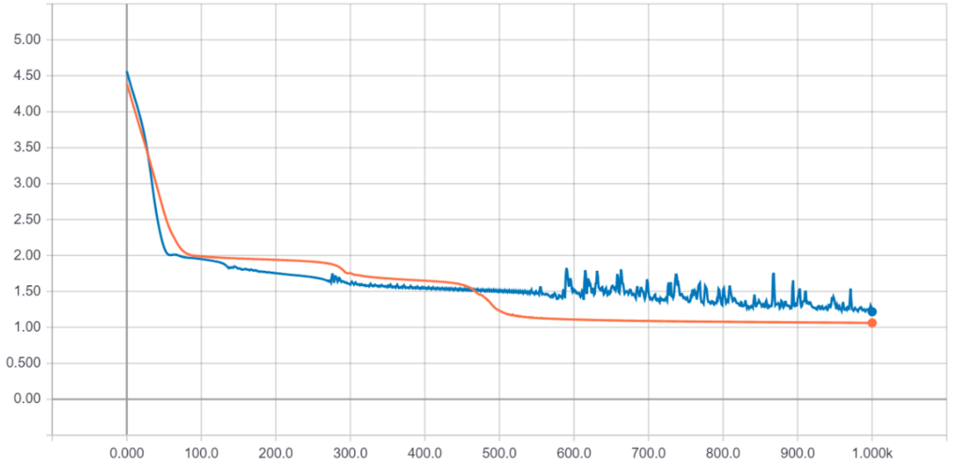

Fig. 4. LSTM (blue) and GRU (orange) loss function plot [14]

However, the most important factor determining the choice of LSTM over GRU is access to many, diverse sources and a much larger database of available knowledge. LSTM is the most commonly used and recommended for this type of project, and it is for this type of network that a large number of guides, open source projects and articles have been created. Thanks to this, solutions to the problems encountered could be found much faster and learn from the experience of a larger group of people who took up the issue in their projects.

3. IMPLEMENTATION STEPS

Realization of the presented issue required both preparation in terms of content and selection of the most optimal network parameters as well as conducting an experiment confirming the effectiveness of the created implementation. For this reason, the entire work was divided into the separated stages.

3.1. Theoretical basis

It was a preparatory stage during which the type of neural network appropriate for the project was selected and the LSTM network operating principles were analyzed. It was also a stage, an important element of which was to get acquainted with the already existing solutions to the problem posed. It is worth mentioning here article named "Deep LSTM with Reinforcement Learning Layer for Financial Trend Prediction in FX High Frequency Trading Systems" [20], which contains a detailed description of the operation of the LSTM network and a proposal to use it to predict currency prices. Another article worth mentioning is "LSTM for time series prediction" [15], which presents the detailed implementation process of the XTB currency value forecast against the US dollar, along with conclusions.

3.2. Creation of the first trial implementation

After getting acquainted with other solutions, selecting the type of network and technology in which the implementation will be made, the first attempts to create it were started. The resulting console application was launched on the Google Colab platform. It is a service offered by Google, allowing for joint work on the code, completing notes, and most importantly, running the code on virtual machines [2]. Thanks to this functionality, it was possible to run multiple instances of the application at the same time, but with different parameters, which significantly accelerated the testing process.

3.3. Results monitoring

To better understand the operation of our network and check its effectiveness, we have implemented monitoring of the training process flow by reporting the size of the loss function in each epoch. We have also added two types of graphs – a graph of the loss function size by epochs (Fig. 5) and a graph comparing the test data with the values predicted during the tests of the model. Thanks to this, we were able to check whether the changes we introduced to individual network parameters and improvements to the program code bring the intended results – they bring us closer to the coverage of the graph of predicted test values with the trends and the course of the function which is representing real values (Fig. 6).

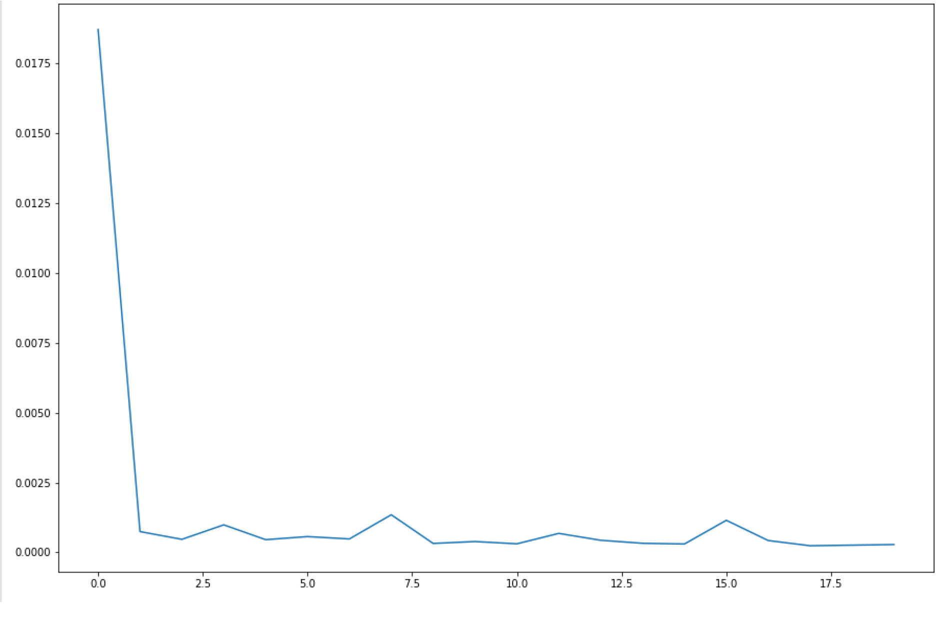

Fig. 5. Sample loss function plot (Source: own study)

Fig. 6. Sample chart that compares real test data with predictions [14]

Ultimately, our network performed best with the following set of parameters:

- Training batch size – 40,
- Test batch size – 1,
- Look back – 10,
- Optimizer – Adam,
- Optimizer learning rate – 0.005,
- Epochs – 70,
- Training data percent – 80%,
- LSTM Units – 128,
- LSTM activation function – ReLU,
- Dense layer units – 1,
- Loss function – Mean Squared Error.

The best effectiveness was considered to be the average result of 96% for ten consecutive launches for Santander Bank Polska.

3.4. Desktop application

After completing the trials and tests and the final selection of the network parameters, we proceeded to implement the desktop application, which will facilitate making predictions and further testing.

Python version 3.6 was used to create the application. The choice made was justified primarily by the popularity of this language in the field of artificial intelligence. In addition, the community around Python has created and made available many libraries for analyzing and processing large amounts of data, creating various types of graphs or for own implementation of neural networks, which is key when writing this type of application. The following libraries were used in the project:

- **Pandas** – open source library for quickly and efficiently manipulating large amounts of data, reading them from CSV format with built-in mechanism for handling missing data [7].
- **NumPy** – an open source library created by the community, used to perform mathematical and logical operations or sorting multidimensional matrices. Additionally, it allows to change the dimension of the matrix. According to the creators, this library is the basic package for scientific computing projects [6].

- **TensorFlow** – created by a team belonging to Google, an open source library used in machine learning, deep neural networks or for the implementation of various types of Machine Learning solutions.
- **Plotly** – open source library, used to create interactive charts, characterized by the possibility of thorough analysis and detailed description of data placed on the chart, which allows to ensure a high level of detail and quality of the created charts.
- **Matplotlib** – an open source, comprehensive library for creating static and animated plots [5].
- **TkInter** – a standard library used to create a graphical user interface used to control the application (GUI) for desktop applications written in Python language.

The application allows predicting three types of time series – the share price of any company listed on WIG20, the price of a unit of any currency in relation to the PLN and the price of a unit of any cryptocurrency in relation to the US dollar. In addition, the application allows for any number of days ahead, any adaptation of network parameters or a constant preview of the learning process by reporting the mean square error in the current epoch. It is final form is shown in Fig. 7.

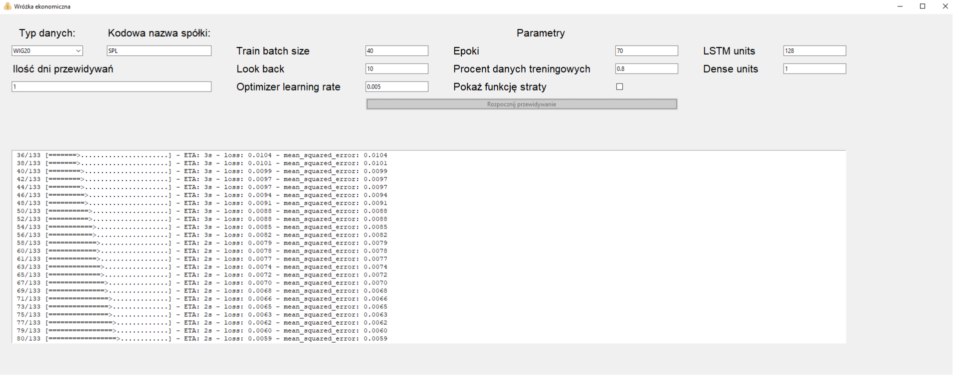

Fig. 7. Desktop application during prediction process (Source: own study)

4. EXPERIMENTS

To check the correctness of the created implementation, its effectiveness and its potential, we decided to conduct two related experiments. The first experiment was to check the effectiveness of the predictions provided by application. The assumptions of the experiment were as follows:

- Forecasts made for 15 consecutive days – from May 25, 2020 to June 8, 2020.
- The predictions were made every day at 22:00 CET (US stock market closing time).
- For each company listed on the WIG20, the value of the company's shares at the close of the stock exchange was predicted.
- The forecasts were made for each company listed on the WIG20.
- For each foreign currency, the unit value in PLN was predicted the next day at 22:00.
- The forecasts were made for the four most important world currencies – the US dollar, euro, British pound and Swiss franc.

- For each cryptocurrency, the unit value in US dollars was predicted the next day at 22:00.
- The predictions were made for six popular (based on market cap) cryptocurrencies – Bitcoin, Litecoin, Ethereum, XRP Coin, Binance Coin and Bitcoin Cash.
- Data on current and historical prices came from two reliable sources. For the Polish stock exchange it was the Stooq portal, for currencies and cryptocurrencies it was a financial portal owned by Yahoo.
- During our forecasts, we used the maximum range of available historical data.

The second experiment was a form of checking whether the created application could replace a real investor. For this purpose, in parallel to the previously described experiment, we invested in cryptocurrencies. The assumptions of this experiment were as follows:

- The initial budget was PLN 20.
- The cryptocurrency was only sold / bought after the predictions for the next day, as described in the previous experiment, had been completed.
- When buying / selling a specific currency, we were guided only by the predictions of our application – we did not interfere with their value. We did not analyze the price charts and information from other sources.
- In case of very optimistic forecasts, it was possible to add a certain amount to the current budget.
- Investments were made using the Revolut application. This platform charges a 1.5% commission on each purchase operation [9], which should also be included in the final results.

An additional goal of the conducted experiments was to show that artificial intelligence can earn on investments made and predict prices with a high probability (oscillating above 95%). Thanks to this, it could replace stock brokers in the future.

4.1. The results of the experiments

After collecting the results from the next 15 days (05/25/2020 – 06/08/2020), the results had to be summed up. To do this objectively, a measure of mean absolute percentage error (MAPE) was used. It informs about the average value of forecast errors for the test period, expressed as a percentage. The MAPE value allows compare the accuracy of the forecasts. Thus, the following formula was chosen to calculate the effectiveness of the predictions (1):

$$E = 1 - \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{1}$$

where the parameters of the formula means:

- E means effectiveness expressed in percentage,
- n means the number of predictions taken for the calculation,
- A_t means real value,
- F_t means forecasted value.

The analysis began with the price forecasts for WIG20 companies. After applying the above-mentioned formula, it turned out that the lowest percentage of the forecasting effectiveness concerned Alior Bank and amounted to 55%. This could be due to the fact that in the last six months, Alior became an unpredictable company, which price dropped from PLN 28 per share to PLN 13 to return to PLN 18 per share. Day by day, the network we created assumed that the company's prices would return to PLN 21, which was probably influenced by the historical, high prices of the company's shares.

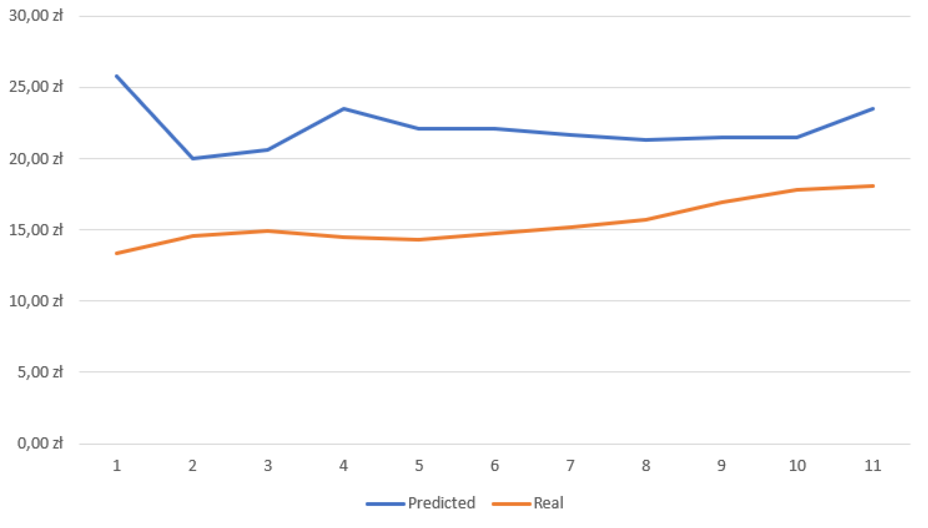

Fig. 8. Alior Bank – plot of predicted stock prices versus real data [19]

In Fig. 8, was shown the discrepancies described previously in the form of a comparison of the predicted stock price chart on the days of the experiment with the actual chart. On the second day of the experiment, there is a clear decline in the projected price relative to the previous day, which can be considered as some value correction. It is also worth noting that on the 10th and 11th day of the experiment, both charts show a clear upward trend. However, the chart of predicted values grows much faster than the real one.

In case of such discrepancies, it is necessary to check the size of the measure, which is the average absolute error expressed in PLN and determined by the formula (2):

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{2}$$

where the parameters of the formula means:

- MAE – mean absolute error in PLN,
- n means the number of predictions taken for the calculation,
- A_t means real value,
- F_t means forecasted value.

For the predicted and real values for Alior Bank, the error value is PLN 6.68, which clearly shows the discrepancy that occurred in the prices of this company.

The highest percentage of the predictions' effectiveness concerned KGHM and Orange Polska and amounted to 98%, which can be considered a very good result. In the case of Orange Polska, a notable feature that could have helped the network achieve such reliable results is the fairly stable share price.

Comparing the chart of projected prices of Orange Polska shares with the chart of actual prices shown in Fig. 9, it should be noted that around the seventh day of the experiment, the network's predictions coincide with the real values, reflecting the trend of price changes, and on the eighth day the predicted value is equal to the actual value. Possible fluctuations are less than 0.10 PLN per share. This fact is confirmed by the calculated value of the mean absolute error, which is exactly 0.10 PLN.

Quite similar conclusions can be presented by looking at the charts for KGHM. Here, the predictions also reflect the real trend, with the difference in prices ranging from 2 to 4 zlotys per share, which is around 2 to 5 percent of the share price.

The lowest value of the mean absolute error was PLN 0.07 and concerned the shares of Tauron Polska Energia (efficiency at the level of 95%). On the other hand, the highest value was PLN 399 for LPP (efficiency at 94%). This shows the dependence of the size of the error on the level of share prices – the lower, the more accurate the predictions of artificial intelligence were. It should be mentioned that LPP is the most valuable company in the WIG20 index. The price per share was PLN 6,930 on July 29, 2020. Tauron is the cheapest company. The price per share was PLN 2.84 on July 29, 2020.

An interesting anomaly that was noticed during the experiment was the lowering of the price of CD Projekt RED shares. Since March, the company has been growing very quickly and has recorded further increases. Most likely, it was this issue that caused our network to be unable to correctly estimate the share price, which resulted in large discrepancies in predictions. For example, on the fourth day of the experiment, the expected price of one share was PLN 263.3, while the real price on that day was PLN 399 – a difference of PLN 135.7 and an efficiency of 66% is a worrying result.

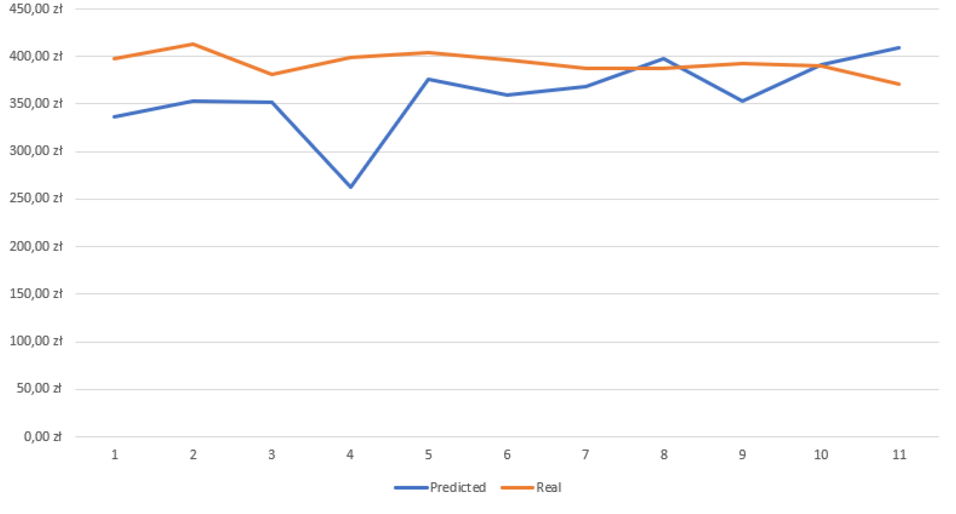

Fig. 9. CD Projekt RED – plot of predicted stock prices versus real data [19]

However, it should also be noted that from the fifth day of the experiment, these values began to equalize, follow trends, and even almost coincide. On the tenth day of the experiment, the effectiveness was 99.8%, and the average finally was 89%.

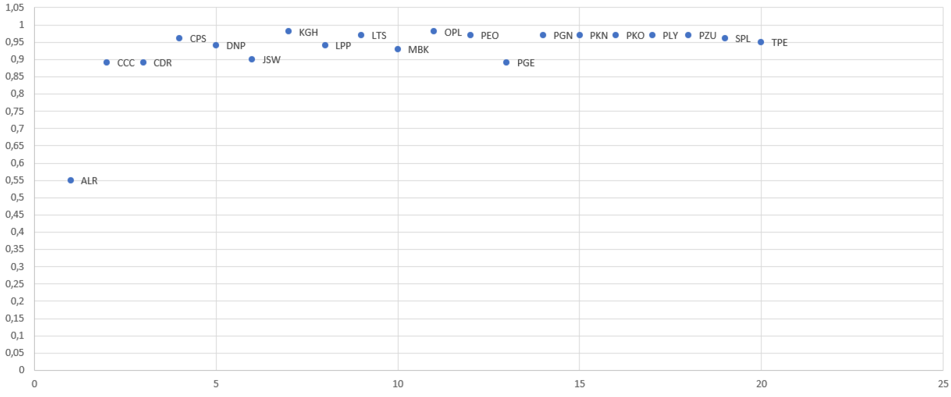

Fig. 10. Distribution of the effectiveness of forecasted prices of WIG20 companies shares (Source: own study)

The final summary of this part of the experiment is the chart of the effectiveness distribution of price predictions for all companies, presented in Fig. 10. It can be noticed that, apart from the abovementioned Alior Bank, the predictions for the remaining 19 companies are 89% and more. For eleven out of twenty companies, the effectiveness was as high as 95%. On the one hand, these results are satisfactory and prove that artificial intelligence can help investors in decisions, and on the other hand, it does not give almost 100% certainty that the predictions will come true.

It is worth mentioning that the average prediction success rate was **92.75%.**

The next step was to analyze the forecasts of currency prices. The results can be summarized by saying that this part of the experiment was characterized by high stability of both actual prices and high forecasts efficiency.

The lowest percentage of the prediction accuracy concerned the US dollar and the Swiss franc – it was 98%. This could be due to the fact that the US dollar recorded a fairly rapid drop in price during the experiment.

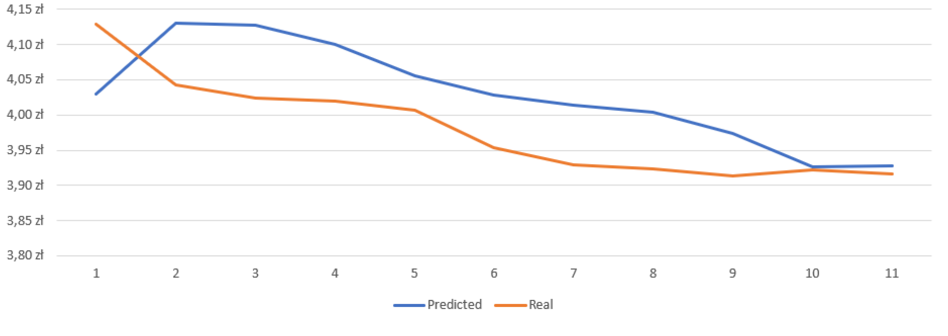

Fig. 11. U.S. Dollar – plot of predicted currency prices versus real data [17]

When analyzing the charts comparison of the predicted prices and the actual, presented in Fig. 11, it is worth noting that even in spite of the drops, the chart representing the expected prices shows a reflection of the actual trend, and the difference in prices was below PLN 0.10.

The highest percentage of predictive success rate was in the euro and British pound – 99%, which can be taken as “**almost certain**”.

This effectiveness can be seen in the individual days of the experiment – on the fifth day of the experiment the predicted price was exactly the same as the real one, on the eighth day the difference was only 0.01 PLN, and on the ninth day 0.02 PLN. Accuracy is also shown by the value of the average absolute error, which is almost 0.04 PLN, exactly 0.034808 PLN. This shows that the application can be successfully used, for example, by people who repay a loan installment in a foreign currency, so that the currency purchase can be made at the best moment with the best price for them.

Ultimately, the average forecast accuracy for currencies was as high as 98.5%. Such a high level of effectiveness is probably due to the stability of the currencies selected for the experiment – behind each of them is a strong economy. In addition, in Poland, we have recently experienced a period of good prosperity and economic growth, thanks to which the domestic currency also does not experience unexpected fluctuations in relation to other world currencies.

An interesting idea to test the operation of the proposed implementation would be to check the effectiveness of forecasts, e.g. for Colombian bolivars. The country is suffering from hyperinflation, which distorts the price stability of the Colombian national currency, and it would be true challenge for AI to predict prices.

The last stage of this experiment remains to be summarized – predicting cryptocurrency prices. We assumed that the cryptocurrency rates, the price of which is driven by the demand of internet users – investors, will be the most difficult to predict. It should be remembered that, compared to national currencies, they do not have a guarantee of stability and the background of large economies. However, looking at effectiveness, our AI has coped with it satisfactorily.

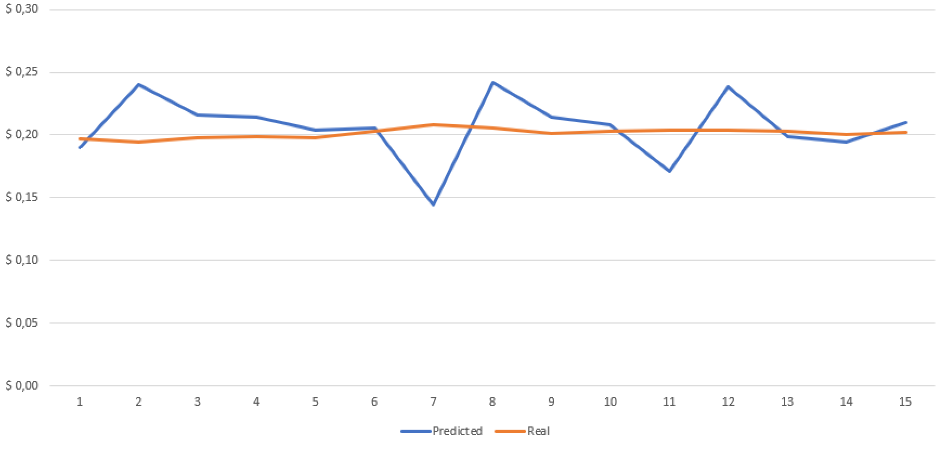

Fig. 12. XRP Coin – plot of predicted cryptocurrency prices versus real data [18]

The lowest percentage of the prediction accuracy concerned the XRP Coin currency and amounted to 90%. This could be due to the fact that the currency is quite stable and its price is around 20 cents, while our AI was able to predict a price decline of up to 14 cents, which was never reflected in reality.

These deviations are best seen in a chart that juxtaposes a chart of predicted and actual prices. It is worth noting, however, that for the last three days the predictions were very accurate, guaranteeing a maximum difference of ± 0.01 $.

The highest percentage of predictive accuracy was for the world's most popular cryptocurrency, Bitcoin, and was 97%. A high result may also be associated with stability, earned position and a wide range of investors.

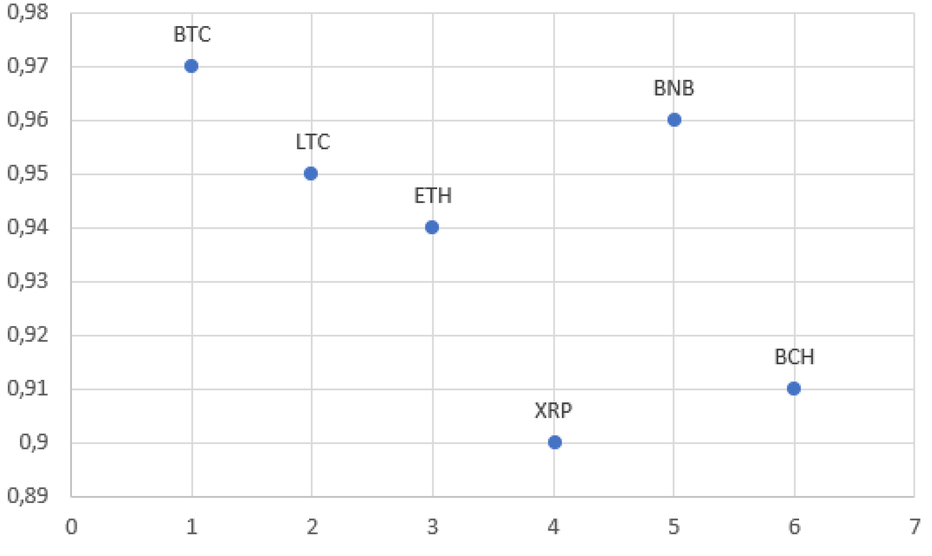

Fig. 13. Distribution of the effectiveness of forecasted prices of cryptocurrencies (Source: own study)

A good summary of this part of the experiment is the fact that the effectiveness of predictions for all cryptocurrencies ranged from 90 to 97 percent, which can be seen in Fig. 13. Thanks to such results, our network can be a real support for investors, being an auxiliary tool in making decisions. However, it is still a tool that does not provide 99/100 percent confidence. In this case, the human being, his knowledge and analysis are required in the investment process too.

After completing the analysis of the first experiment, it is time to summarize the investment experiment, which was to show that our AI can replace a real investor or at least become a real help for those working in this profession. Summarizing the results of investments driven by the predictions of artificial intelligence, it can be stated that:

- the total budget used in the experiment was 45 PLN,
- the total commission for the purchase of cryptocurrencies was 3.34 PLN (7.42% of the total budget),
- the budget on the last day of the experiment was 44.95 PLN (99.89% of the initial budget),
- excluding the commission, the final balance sheet was +3.29 PLN (7.31% of the total budget). This means that the entire experiment ended with a slight profit.

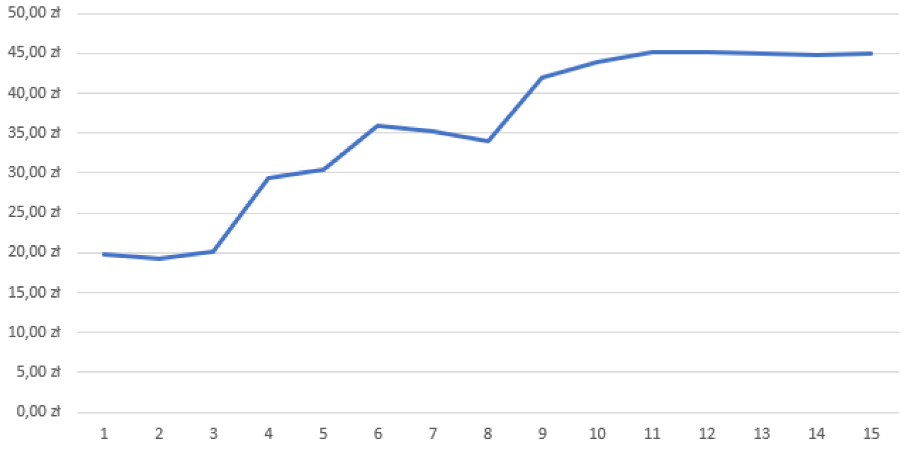

Fig. 14. Plot shows budget changes during the experiment (Source: own study)

To better understand how your experiment works, it is necessary to analyze daily budget chart. At the beginning, in order to eliminate unnatural increases, it should be noted when the budget has been increased by additional funds.

- The first deposit, in the amount of 10 PLN, was made on the third day.
- The second deposit, in the amount of 5 PLN, was made on the fifth day.
- The third deposit, in the amount of 10 PLN, was made on the eighth day.

In addition, the batch size parameter was modified during the experiment to check how it would affect the results. The change took place on the seventh day of the experiment. The value of the parameter was reduced to 30, however, as can be seen in the chart, it resulted in the greatest decrease, therefore the previous parameter value of 40 was immediately restored.

By analyzing the chart itself, you can see that from the tenth day of the experiment there is a slow and gentle increase in the budget. This period can be associated with the effectiveness of predictions, which for all the cryptocurrencies we owned at that time (BTC, ETH, LTC) was over 95%.

5. COMPARISON RESULTS OF THE EXPERIMENT PERFORMED WITH OTHERS AVAILABLE SOLUTIONS

Many external sites allow to purchase software similar to that created by the authors. There are also open-source hobby projects on the Internet dealing with price predictions. By averaging all the results of the prediction of the application, we get an efficiency of about 93% to as much as 98.5% in the case of currencies. With this in mind – the application has some of the best results compared to other solutions offered on the market. Below is a description of other solutions based on publicly available data:

- **predictioncoins.com**
  A website specializing in cryptocurrency investments. They offer an application functionally similar to the one that was proposed during the research. The accuracy of the results of the offered application oscillates around 85%. Unfortunately, the producer does not state in the public data what type of network was used here for predictions.

- **bitcurate.com**
  A website that offers a application with prediction accuracy of up to 80%. In addition, the user has a lot of additional information at his disposal, such as charts presenting company prices and news from the world, which allows easier decision-making and increasing the above-mentioned. accuracy through human-system interaction.
- **cryptics.tech**
  Declared predictions with an accuracy of 70%.
- **intupercepts.com**
  This service offers the most likely results with an accuracy of as much as 90%. It is one of the most promising applications on the market. Additionally, the manufacturer offers to rewrite the program into any language chosen by the client.
- **Open source projects hosted on Github**
  - https://github.com/Btsan/Crypto-Price-Prediction – author made predictions based on data from 365 days. The declared result was an efficiency of 55%.
  - https://github.com/huseinzol05/Stock-Prediction-Models – declared efficiency at the level of 95.693% for the LSTM network

In each of the above solutions, an LSTM neural network was used. The conclusion from these research is also the fact that most of the proposed solutions use this type of network. The data was searched on the basis of popular websites like GitHub or GitLab and applications.

5.1. Development and improvement of the existing project

5.1.1. Monitor and track events in politics, society and economy

First of all, the most important disadvantage of our application was the lack of resistance to unexpected events. We relied solely on the price figures at the close of the stock exchange, not taking into account global events that also affect investor sentiment. The direction of application development in this field could be the implementation of the analysis of the largest economic or news websites in terms of information on companies. Then, this information would be classified in terms of emotions emanating from it, and this element would also be taken into account when predicting prices for the next day or during the current trading session. This technique is called ”sentiment analysis”.

5.1.2. Implementation of the analysis of financial markets technical indicators

The implementation of the analysis of technical indicators could significantly support the accuracy of predictions. They would allow for the correction of forecasts taking into account the probability of ups and downs and the size of price fluctuations. This probability could be determined by one of several indicators used in the financial field. Examples of indicators are:

- **Relative strength index** which provides signals that tell investors to buy when the security or currency is oversold and to sell when it is overbought [10]. These are signals that hints the investor to sell or buy stocks. Appropriate values adopted by this indicator also allow to determine the probability of a trend reversal into a downtrend or an upward trend.

- **Moving Average Convergence/Divergence** (MACD) which is designed to reveal changes in the strength, direction, momentum, and duration of a trend in a stock's price [3]. This indicator is composed from two lines – MACD and signal line. The crossing of the two lines is an indicator of the increased probability of a trend change and a signal for the investor to sell or buy the stock.
- **Average True Range** (ATR) which was created to allow traders to more accurately measure the daily volatility of an asset. A stock experiencing a high level of volatility has a higher ATR, and a low volatility stock has a lower ATR [8]. This indicator allows the risk of fluctuations in share prices to be assessed when making a purchase decision.

5.1.3. Include market anomalies

In order to better adapt to the behavior of investors and improve the application's resistance to various types of unforeseen behavior, one should also take into account market anomalies, i.e. "a situation enabling the achievement of positive, above-average rates of return" [16]. Examples of anomalies are:

- **"Monday effect"** characterized by negative rates of return on that day of the week [21].
- **"Friday effect"**, characterized by a positive average rate of return, and at the same time the highest value of this measure compared to other days of the week [21].
- **"sell in may and go away"** is an effect related to the holiday period, before which investors get rid of part of their portfolio, which is associated with drops [21].
- **"January effect"**, which consists of several elements, i.e. supplementing the composition of investment portfolios by investors and repurchasing shares sold in December of the previous year [12].
- **"Santa Claus rally"** or **"December effect"** occurs when investors buy shares in early December, with hope to sell them back at a profit in January using the "January effect" [4].

5.1.4. Use of supervised training

This method of teaching would be much longer, but thanks to the supervisor supporting the learning process, it would be possible to obtain results oscillating around 99–100% efficiency. The whole process would be to support the application by making adjustments to neuron weights and telling it what output should be. Of course, this is only a hypothesis, but the fact is that in our opinion it would definitely improve the final result.

6. DIFFICULTIES AND CONCLUSIONS

A few days after project began, the process of "freezing" national economies for fear of the coronavirus began around the world. Economists threatened a huge crisis. There were many phenomena on the Polish stock exchange that were impossible to predict only through the analysis of previous data – the shares of one of the largest state-owned energy companies – Tauron Polska Energia SA fell below PLN 1, and CD Projekt RED (game producer) became the strongest company on WIG20. Additionally, during the experiment, the US dollar suffered an unexpected decline due to race unrest.

However, these are not the only unforeseeable factors that influence this type of project. In the era of easy access to information – one person's impulse is literally enough for the entire stock exchange to fall into recession. A clear example would be the following hypothetical scenario: An article titled "The Dollar Is About to Collapse?" has been published. In this article the author discusses his thoughts about worldwide position of dollar for the next 100 years (the US dollar became a full-fledged means of payment in 1792 [16]), which on the currency's life timeline means that 100 years may actually be close to the event). At this point, the person with the money invested in this currency is selling it because he has not noticed that this is just a logical analysis of the author's 100-year currency future hypothesis. It turns out that this person had a fairly large sum of money that he exchanged instant. Other people seeing instant value drop are also selling dollars to earn as much as possible before the decline. At this point, a queue of people massively selling the currency begins, panic breaks out in the market and the dollar value plummets. It also entails a recession in the American economy and declines in world stock markets. After all, the entire world economy is a system of interconnected vessels,. Therefore, the decision of one person significantly influenced one of the world's strongest currencies.

This is only a simple and abstract analogy with a very low probability of an event, but it shows the scale of the problem best. On the basis of data from even over a dozen previous years, no one is able to predict a hypothetical panic on the stock exchange caused, for example, by unauthorized access to the collected funds. A fairly realistic example is the hacking attack and the problems of the Mt. Gox from February 2014, as a result of which the Bitcoin cryptocurrency price dropped by as much as 36% – from $ 727 on February 7, 2014 to $ 466 at the end of March [1].

Here are some conclusions that could be made while learning a neural network based solely on the analysis of the daily, monthly and yearly graphs of any course:

- The project is based on the above-mentioned the input data will never take into account external factors such as future pandemics, protests, wars etc.
- Even the best optimized network will not be 100% accurate – it is the decisions of individual people and the situation in the world that shape the stock market. In the case of cryptocurrencies, these are not only politicians or brokerage houses, but also ordinary users who have invested in it.
- The received predictions will only be a hint that a investor can follow when making a decision to invest. Equally important is the simultaneous analysis of other factors influencing the stock market.

REFERENCES

[1] Ashton M., 2016. What's Wrong with Money?: The Biggest Bubble of All. John Wiley & Sons.
[2] Based on https://colab.research.google.com/notebooks/intro.ipynb.
[3] Based on https://en.wikipedia.org/wiki/MACD.
[4] Based on https://histmag.org/Dolar-historia-zielonego-banknotu-15826/.
[5] Based on https://matplotlib.org.
[6] Based on https://numpy.org/doc/stable/user/whatisnumpy.html.
[7] Based on https://pandas.pydata.org/about/.
[8] Based on https://www.investopedia.com/terms/a/atr.asp.
[9] Based on https://www.revolut.com/legal/fees, option for premium customer.
[10] Deng, S., Sakurai, A., 2013. Foreign Exchange Trading Rules Using a Single Technical Indicator from Multiple Timeframes. 27th International Conference on Advanced Information Networking and Applications Workshops, 207–212.
[11] Hochreiter S., Schmidhuber J., 1997. Long short-term memory. Neural Computation 9(8), 1735–1780, Massachusetts Institute of Technology Press.

[12] Lewandowska M., 2017. Efekt stycznia i efekt grudnia na Giełdzie Papierów Wartościowych w Warszawie. Journal of Capital Market and Behavioral Finance 1(5), 17–28.
[13] Michael P., 2018. Retrieved from https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-
[14] Muccino E., 6 march 2019, LSTM vs GRU: Experimental Comparison. Retrieved from https://medium.com/mindboard/lstm-vs-gru-experimental-comparison-955820c21e8b.
[15] Orac R., 28 september 2019, LSTM for time series prediction. Retrieved from https://towardsdatascience.com/lstm-for-time-series-prediction-de8aeb26f2ca.
[16] Peters E., 1997. Teoria chaosu a rynki kapitałowe WIG-Press Warszawa.
[17] Real currency prices based on data provided by Finance Yahoo, https://finance.yahoo.com/currencies.
[18] Real currency prices based on data provided by Finance Yahoo, https://finance.yahoo.com/cryptocurrencies.
[19] Real WIG20 companies shares prices based on data provided by the Stooq, https://stooq.pl/t/?i=532.
[20] Rundo F., 2019. Deep LSTM with Reinforcement Learning Layer for Financial Trend Prediction in FX High Frequency Trading Systems. Applied Sciences 9(20), 4460.
[21] Zawadzki A., Troska D., Domańska M., 2017. Anomalie kalendarzowe na Giełdzie Papierów Wartościowych w Warszawie. Journal of Capital Market and Behavioral Finance 1(5), 7–15.

PROGNOZOWANIE PRZYSZŁYCH WARTOŚCI SZEREGÓW CZASOWYCH Z WYKORZYSTANIEM SIECI LSTM NA PRZYKŁADZIE KURSÓW WALUT I SPÓŁEK WIG20

Streszczenie

W artykule przedstawiono porównanie sieci RNN, GRU i LSTM w przewidywaniu przyszłych wartości szeregów czasowych na przykładzie walut i spółek giełdowych. Przedstawiono również etapy tworzenia aplikacji będącej realizacją analizowanego zagadnienia – dobór sieci, technologii, dobór optymalnych parametrów sieci. Dodatkowo omówiono dwa przeprowadzone eksperymenty. Pierwszym było przewidywanie kolejnych wartości spółek z WIG20, kursów walut i kryptowalut. Drugi opierał się na inwestycjach w kryptowaluty, kierując się wyłącznie przewidywaniami sztucznej inteligencji. Miało to na celu sprawdzenie, czy inwestowanie na podstawie przewidywania takiego programu pozwala na efektywne zarobki. Omówienie wyników eksperymentu obejmuje analizę różnych ciekawych zjawisk, które wystąpiły w czasie jego trwania oraz kompleksowe przedstawienie relatywnie wysokiej skuteczności proponowanego rozwiązania wraz z wszelkiego rodzaju wykresami i porównaniami z rzeczywistymi danymi. Analizowano również trudności, które wystąpiły podczas eksperymentów, takie jak koronawirus, wydarzenia społeczno-gospodarcze czy zamieszki w USA. Na koniec zaproponowano elementy, które należałoby ulepszyć lub uwzględnić w przyszłych wersjach rozwiązania, uwzględniając wydarzenia na świecie, anomalie rynkowe oraz wykorzystanie uczenia się nadzorowanego.

Słowa kluczowe: rekurencyjna sieć neuronowa (RNN), blok rekurencyjny (GRU), pamięć długookresowa (LSTM)