ATLANTIS PRESS

Advances in Intelligent Systems Research, volume 174

Proceedings of the 8th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2020)

Using a Combination of Recurrent and Convolutional Neural Networks to Forecast the Direction of Financial Instrument Price Movement

V. A. Melnikov
Institute of Information Technologies
Chelyabinsk State University
Chelyabinsk, Russia
protei300@gmail.com

N. D. Kharchenko*
Institute of Information Technologies
Chelyabinsk State University
Chelyabinsk, Russia
wrk.altnar@gmail.com

***Abstract*—Securities market forecasting has long been of interest to analysts and mathematicians due to the obvious opportunity to monetize the research if it proves to be successful. The work of these researchers has led to the creation of various trading algorithms; however, their effectiveness has not yet been proven. With the development of computing technologies that allow implementing complex mathematical machine learning systems, the attention to this direction has increased considerably, in particular because of the introduction of neural networks. The present paper focuses on describing the initial data (pairs of price and the number of transactions available at this price) and the process of data collection and preparation for the neural network training. Moreover, the reasons for choosing the combination of recurrent and convolutional neural networks and its scheme are given, and the training results and insights are presented.**

Keywords—recurrent neural networks, convolutional neural networks, open interest, forecasting, financial market

I. INTRODUCTION

Our previous research has shown that the forecasts of financial instrument movement based solely on the previous data on the instrument price are inconsistent [1]. In the present paper, we study the possibility of forecasting the financial market, using the open interest data. This data source is actively used by professional traders and demonstrates the market participants' interest in certain cost positions of a financial instrument.

Open interest can be represented as two sets of rows. Each row stores the price value and the number of transactions available at that price (trading volume): the first (upper) set of rows contains information on purchase transactions, and the second (lower) set contains information on sales transactions. Due to the large number of trading participants, the open interest can consist of hundreds of rows; however, the further they are from the center, the less they influence the market. Therefore, it was decided to use only 20 rows closest to the center: 10 from the lower part of the set and 10 from the upper part.

A neural network based on the combination of convolutional and recurrent layers was chosen as an algorithm for analyzing and forecasting the financial market, using the received data. Such layers have proven effective in computer vision, text analysis, and speech recognition [2-5].

Convolutional layers of a neural network help to reduce the number of parameters and to represent the results in a more abstract form. Therefore, this allows building more complex neural network models and provides the increased effectiveness of the layers that follow the convolutional one [6, 7].

Since time series were used in the present research, the second component of the model was the long short-term memory layers – a type of recurrent neural networks. The main idea of a recurrent network is that it is able to compare the examples from different points in time and to establish a pattern [8, 9]. Nevertheless, as their learning process showed, when it was required to find a dependence between distant events, it was almost impossible to fit the necessary parameters [10].

Although, when using the back-propagation of error algorithm, a modification with a long short-term memory can circumvent this restriction due to the gradient descent mechanism [11, 12].

II. DATA PREPARATION

In preparation for the experiment, open interest data on Light Sweet Crude Oil from the Chicago Mercantile Exchange were collected. The data were recorded with a frequency of 1 second during the bidding from October 3, 2019 to October 9, 2019. The total data amount was 290,000, and for training, random sections of 170,000 rows were selected. The data for sale price and its volume respectively were presented as $p^i_b(t)$ and $V^i_b(t)$, and, similarly, the data for purchase price and its volume respectively were presented as $p^i_a(t)$ and $V^i_a(t)$, where i is the level of the open interest $i \in [1,10]$. Thus, the open interest vector was presented as the following (1):

$$x = \{p^i{}_a(t), V^i{}_a(t), p^i{}_b(t), V^i{}_b(t)\}_{i=1}^{n=10}. \quad (1)$$

Data labeling allowed us to obtain information on the future direction of the price movement, and to do this, average price for a financial instrument at the point in time had to be calculated by comparing two prices closest to the center (2):

$$p_t = \frac{p^1{}_a(t) + p^1{}_b(t)}{2}. \quad (2)$$

Since financial instrument prices can change significantly at different points in time, the most rational way to determine the direction of their movement was to use the percentage change between prices at different points in time. Analysis of other papers on the related topic permitted us to find the following technique for determining the price movement (3, 4):

$$m_-(t) = \frac{1}{k}\sum_{i=0}^{k} p_{t-i}\,, \tag{3}$$

$$m_+(t) = \frac{1}{k}\sum_{i=0}^{k} p_{t+i}\,, \tag{4}$$

where m is the movement and k is the forecast period [13]. Therefore, the percentage change of the average price over the l_t period was (5):

$$l_t = \frac{m_+(t) - m_-(t)}{m_-(t)}. \tag{5}$$

As a target feature, we used the percentage change of the financial instrument price by a certain value, where 0 is the absence of the sufficient price fluctuation, 1 is the increase, and 2 is the decrease. This rule can be written as follows (6):

$$\begin{cases} l_t > \alpha, 1 \\ \alpha \le l_t \le \alpha, 0 \\ l_t < -\alpha, 2 \end{cases}, \tag{6}$$

where a is the required change value.

Within the framework of the present study, a equaled to 0.0003% of the transaction value.

III. Neural Network

To develop a prediction model, a system based on using the layers of the one-dimensional convolutional neural network (CNN) and the layers of the recurrent neural network with long short-term memory (LSTM) was used [14, 15].

Convolutional layers were used for automatic feature extraction. The reason for choosing a one-dimensional CNN was that the full information on the open interest condition was stored in each separate row [16].

LSTM is the most effective tool for working with time series due to the fact that there is a negative connection that enables saving information [17]. Several connected LSTM layers were also used at once, and each of them displayed a sequence of vectors that would be used as input data to the subsequent LSTM layers [18]. This hierarchical system permitted us to display the time series data more comprehensively by transforming the information into different scales (Fig. 1) [19, 20].

IV. Training Results

As a result of a repeated neural network training, the following data in the training set were obtained (Table I):

TABLE I.

	Precision	Recall	F1 Score
1	62	29	40
0	71	79	75
2	77	28	41

From the obtained data, it can be indicated that class 1 and class 2, which illustrate that the price has increased or decreased by more than the threshold value a, are predicted in a way that in most cases the final forecast permits us to make the best transaction (the Precision value reflects what percentage of all positive predictions have been correct); however, only a small part of all possible profitable transactions will be made (the Recall parameter shows what percentage of all positive outcomes have been predicted). It can be concluded that the model is only capable to forecast properly the average price stability, which is not very useful. Nevertheless, if the scope of the experiment is taken into account, the high Precision level of all three classes can provide a high percentage of profitable transactions even though they will not be made often.

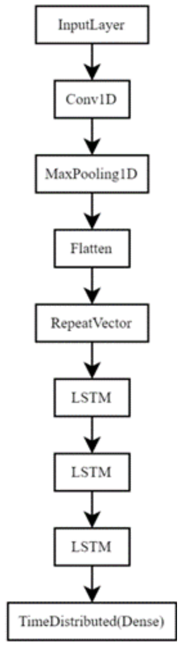

Fig. 1. Neural network model.

V. Conclusion

The findings of this study confirm the possibility of using a combination of recurrent and convolutional neural networks to forecast the direction of financial instrument price movement. To further the research, we plan to use the described model or its improved versions to develop a profitable automated trading system.

References

[1] V. A. Melnikov and N. D. Kharchenko, “Evaluation of the time series predictability in the foreign exchange market using machine learning methods” (in Russian), in Inf. Technologies and Systems 2019, Khanty-Mansiysk, Russia, 2019, pp. 74–78.

[2] J. Koo, D. Klabjan, and J. Utke, “Combined convolutional and recurrent neural networks for hierarchical classification of images,” 2018, arXiv:1809.09574.

[3] X. Wang, W Jiang, and Z. Luo, “Combination of convolutional and recurrent neural network for sentiment analysis of short texts,” in Proc. COLING 2016: 26th Int. Conf. Comput. Linguistics, Osaka, Japan, 2016, pp. 2428–2429.

[4] N. Takahashi, N. Goswami, and Y. Mitsufuji, “MMDenseLSTM: An efficient combination of convolutional and recurrent neural networks for audio source separation,” in 2018 16th Int. Workshop on Acoustic Signal Enhancement, Tokyo, Japan, 2018, pp. 106–110.

[5] G. Liang, H. Hong, W. Xie, and L. Zheng, “Combining convolutional neural network with recursive neural network for blood cell image classification,” IEEE Access, vol. 6, pp. 36188–36197, 2018.

[6] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in 2017 Int. Conf. Engineering and Technol., Antalya, Turkey, 2017, pp. 1–6.

[7] K. O'Shea and R. Nash, “An introduction to convolutional neural networks,” 2015, arXiv:1511.08458.

[8] S. Dupond, “A thorough review on the current advance of neural network structures,” Annual Reviews in Control, pp. 200–230, 2019.

[9] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” Nature, vol. 323, no. 6088, pp. 533–536, 1986.

[10] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” IEEE Transactions on Neural Networks, vol. 5, no. 2, pp. 157–166, 1994.

[11] F. A. Gers, N. N. Schraudolph, and J. Schmidhuber, “Learning precise timing with LSTM recurrent networks,” The Journal of Machine Learning Research, vol. 3, pp. 115–143, 2002.

[12] F. A. Gers, J. Schmidhuber, and F. Cummins, “Learning to forget: Continual prediction with LSTM,” in 9th Int. Conf. Artificial Neural Networks ICANN 99, Edinburgh, UK, 1999, vol. 2, pp. 850–855.

[13] A. Tsantekidis, N. Passalis, A. Tefas, J. Kanniainen, M. Gabbouj, and A. Iosifidis, “Forecasting stock prices from the limit order book using convolutional neural networks,” in Conf. Business Informatics 2017, Thessaloniki, Greece, 2017, vol. 1, pp. 7–12.

[14] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[15] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” The Journal of Machine Learning Research, vol. 15, no. 1, pp. 1930–1931, 2014.

[16] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, “1D Convolutional Neural Networks and Applications: A survey,” pp. 6–9, 2019.

[17] S. Siami-Namini, N. Tavakoli, and A. S. Namin, “A comparison of ARIMA and LSTM in forecasting time series,” in 2018 17th IEEE Int. Conf. Machine Learning and Applications (ICMLA), Orlando, Florida, USA, 2018, pp. 1394–1401.

[18] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, “Single layer & multi-layer long short-term memory (LSTM) model with intermediate variables for weather forecasting,” in 3rd Int. Conf. on Comput. Science and Comput. Intell., 2018, pp. 90–98.

[19] F. Karim, S. Majumdar, and H. Darabi, “Insights into LSTM fully convolutional networks for time series classification,” IEEE Access, vol. 7, pp. 67718–67725, 2019.

[20] Q. Yin, R. Zhang, and X. Shao, “CNN and RNN mixed model for image classification,” MATEC Web of Conferences, vol. 277, 2019.