https://amf.ui.ac.ir

Journal of Asset Management and Financing
E-ISSN: 2383-1189
Vol.10, Issue 3, No. 38, Autumn 2022, p 1-20
Received: 28.06.2021 Accepted: 07.03.2022

Research Paper

Application of Deep Learning Architectures in Stock Price Forecasting: A Convolutional Neural Network Approach

Amir Sharif Far
Ph. D. Student, Department of Financial Management, Faculty of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran. Iran.
am.sharifar@gmail.com

Maryam Khalili Araghi *
Assistant Professor, Department of Business Management, Faculty of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran. Iran.
m.khaliliaraghi@srbiau.ac.ir

Iman Raeesi Vanani
Associate Professor, Department of Industrial Management. Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran. Iran.
imanraeesi@atu.ac.ir

Mirfeiz Fallahshams
Associate Professor, Department of Financial Management, Faculty of Management, Central Tehran Branch, Islamic Azad University, Tehran,
Iran.fallahshams@gmail.com

Abstract
Algorithms based on a Convolutional Neural Network (CNN), which is a branch of Deep Learning (DL), have seen significant progress in picture and video analyses in recent years. Success of these new models has led to widespread use of them in various fields, including text mining and time series data. DL is part of a broader family of machine learning methods that attempts to model high-level concepts using learning at multiple levels and layers and extract features of higher levels from the raw input. This survey investigated the abilities of different CNN architectures to predict stock prices. Upon running the model with various architectures and parameters for the stock price of Esfahan Steel Company, the results showed that a CNN with max-pooling layers (a combination of Batch size=64, filters=256, and ReLU Activation Function) and Mean Absolute Percentage Error (MAPE) of 1.79% and Normalized Root Mean Square Error (NRMSE) of 2.71% had a higher prediction accuracy than other CNN architectures and Recurrent Neural Network (RNN).
Keywords: Stock price prediction, Machine learning, Deep learning (DL), Convolutional neural network (CNN), Recurrent neural network (RNN)

Introduction

Among the various deep learning techniques that have many applications in different sciences, specific algorithms like Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN) have been used by researchers due to their characteristics of financial time series (Sezer, Gudelek, & Ozbayoglu, 2020). CNN is a feed-forward Artificial Neural Network (ANN) that takes its inputs as 2-D matrices. Unlike a fully connected neural network like Multi-Layer Perception (MLP) neural network, the locality of data within its input vector (or matrix) is important (Sezer & Ozbayoglu, 2018).

CNN has different architectures. Usually one specific architecture is focused on in each study conducted in this field. In this study, however, the architectures used in various studies were surveyed in the first level and each selected architecture was optimized by using different parameters in the second level. Finally, the best performances of the architectures with various parameters were compared to choose the optimized model. The effective studies in model development are shown in Table 1.

*Corresponding author
Sharif Far, A., Khalili Araghi, M., Raeesi Vanani, I. & Fallahshams, M. (2021). Application of Deep Learning Architectures in Stock Price Forecasting: A Convolutional Neural Network Approach. *Journal of Asset Management and Financing*, 10(3), 1-20.

2383-1189 / © 2022

This is an open access article under the BY-NC-ND /4.0/ License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

http://dx.doi.org/10.22108/AMF.2022.129205.1673 https://dorl.net/dor/20.1001.1.23831189.1401.10.3.1.9

Table (1) Effective studies in model development

Art.	Method	Dataset	Feature Set	Horizon
Livieris, E. Pintelas, & P. Pintelas (2020)	Using two convolutional layers with different filters	Gold	Price data	1 day
Gao, Zhang, &Yang (2020)	Simple CNN	S & P500 CSI300 Nikkei225	Price data, volume, technical indicators	1 day
	CNN with a dropout layer			
Gudelek, Boluk, & Ozbayoglu (2017)	CNN with dropout and max-pooling layers	ETF	Price data, technical indicators	1 day
Ji, Zou, He, & Zhu (2019)	CNN with a max-pooling layer	Future carbon price	Price data	7 days
Li & Dai (2020)	CNN with a max-pooling layer	Bitcoin	Price data	1 day

Method and Data

Based on the previous studies on CNN application, three different architectures of CNN were investigated as shown in Figure 1.

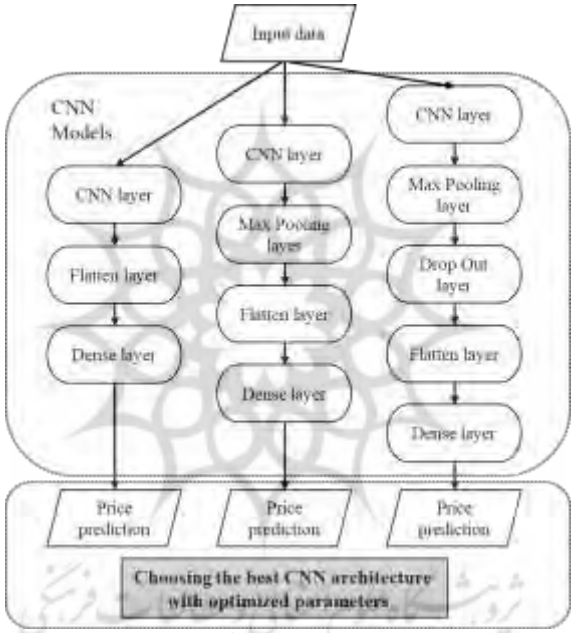


Figure (1) The process of choosing an optimal CNN algorithm

For selecting the best CNN architecture, all the three models were surveyed with various parameters. It is worth noting that the parameters that affected CNN included items like number of filters in the CNN layer, Batch size, and Activation Function. In this study, the data obtained from Esfahan Steel Company during the period of 2018-2021were used. The input data consisted of two categories, including price data (Open, High, Low, Close, and Volume) and technical indicators based on the surveys of Kara et al. (2011) and Patel et al., (2015). Python 3.8 with Keras Library was used to execute the model. In this study, the dataset was divided into a training set and a testing set, which covered about the first 80% and last 20% of the raw dataset, respectively.

Findings

Comparison of the three defined architectures with various parameters led to the optimized model. It should be noted that the selected model was the result of running it 54 times with different layers and parameters. In this study, the two performance measures of Mean Absolute Percentage Error (MAPE) and Normalized Root Mean Square Error (NRMSE) were selected to evaluate the predictive power of our proposed models. In Table 2, the errors of the best performances of each of the three architectures with different parameters and the RNN model (another DL model) were compared to choose the optimized model. Based on the results, the accuracy of the best performance of the second CNN architecture was higher than those of the others.

Table (2) Comparison of the errors of the selected models

Error / Method	MAPE	NRMSE
RNN	2.46%	2.79%

Best performance of the first CNN architecture	2.13%	3.09%
Best performance of the second CNN architecture	1.79%	2.71%
Best performance of the third CNN architecture	2.18%	3.26%

Conclusion and discussion

In this paper, the predictive powers of the various architectures of CNN models were investigated. The results demonstrated that the best performance of the second CNN architecture with the Max-Pooling layer and combination of Batch size of 64, filter of 256, and ReLU Activation Function and MAPE and NRMSE errors of 1.79 and 2.71%, respectively, provided higher prediction accuracy than other CNN and RNN architectures. The outcome of this survey was supported by research of Ji et al. (2019) on Carbon future price forecasting and that of Li & Dai (2020) on Bitcoin price forecasting by using a CNN model with convolutional and Max-Pooling layers.

However, Gao et al. (2020) proposed a convolutional layer with a dropout layer and Gudelek et al. (2017) used a convolutional layer with dropout and Max-Pooling layers for predicting ETF prices. Their results were not confirmed by this paper since using a convolutional layer with a Max-Pooling layer had a better performance than other architectures.

References

Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, 519, 127-139. https://doi.org/10.1016/j.physa.2018.11.061

Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of china stock market. *In 2015 IEEE International Conference on Big Data (Big Data)*, 2823-2824. https://doi.org/10.1109/BigData.2015.7364089

Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 83,187-205. https://doi.org/10.1016/j.eswa.2017.04.030

Dami, S., & Khodabaksh, J. F. (2017). Predicting the price of stock market using deep learning. *National Conference on Modern Knowledge and Technology in Engineering in the Technology Era.* Tehran. (In Persian)

Das, S. R., Mokashi, K., & Culkin, R. (2018). Are markets truly efficient? Experiments using deep learning algorithms for market movement prediction. *Algorithms*, 11(9), 138-157. https://doi.org/10.3390/a11090138

Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing*, 7(3–4), 197-387. http://dx.doi.org/10.1561/2000000039

Dezsi, E., & Nistor, I. A. (2016). Can deep machine learning outsmart the market? A comparison between econometric modelling and long- short term memory. *Romanian Economic Business Review*, 11(4.1), 54-73.

Gao, P., Zhang, R., & Yang, X. (2020). The application of stock index price prediction with neural network. *Mathematical and Computational Applications*, 25(3), 53-69. https://doi.org/10.3390/mca25030053

Gordon, J. A., & Francis, J. C. (1986). *Portfolio Analysis, (3th. Ed.).* California: Prentice-Hall.

Gudelek, M. U., Boluk, S. A., & Ozbayoglu, A. M. (2017). A deep learning based stock trading model with 2-D CNN trend detection. *In 2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1-8.

Gunduz, H., Yaslan, Y., & Cataltepe, Z. (2017). Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations. *Knowledge-Based Systems*, 137, 138-148. https://doi.org/10.1016/j.knosys.2017.09.023

Hiransha, M., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2018). Nse stock market prediction using deep-learning models. *Procedia Computer Science*, 132, 1351-1362. https://doi.org/10.1016/j.procs.2018.05.050

Hoseinzade, E., & Haratizadeh, S. (2019). CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, 129, 273-285. https://doi.org/10.1016/j.eswa.2019.03.029

Ji, L., Zou, Y., He, K., & Zhu, B. (2019). Carbon futures price forecasting based with ARIMA-CNN-LSTM model. *Procedia Computer Science*, 162, 33-38. https://doi.org/10.1016/j.procs.2019.11.254

Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert systems with Application*, 38 (5), 5311-5319. https://doi.org/10.1016/j.eswa.2010.10.027

Khare, K., Darekar, O., Gupta, P., & Attar, V. Z. (2017). Short term stock price prediction using deep learning. *In 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 482-486. https://www.doi.org/10.1109/RTEICT.2017.8256643

Kiyani Mavi, R., & Sayadi Nik, K. (2015). Using different learning algorithms in the stock price prediction by using neural networks. *Journal of Development and Evolution Management*, 1393, 75-81. (In Persian)

Lee, S. I., & Yoo, S. J. (2020). Threshold-based portfolio: The role of the threshold and its applications. *The Journal of Supercomputing*, 76(10), 8040-8057. https://doi.org/10.1007/s11227-018-2577-1

Li, Y., & Dai, W. (2020). Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model. *The Journal of Engineering*, 2020(13), 344-347. https://doi.org/10.1049/joe.2019.1203

Liu, S., Zhang, C., & Ma, J. (2017). CNN-LSTM neural network model for quantitative strategy analysis in stock markets. *International Conference on Neural Information Processing*, Springer, Cham. https://doi.org/10.1007/978-3-319-70096-0_21

Livieris, I. E., Kiriakidou, N., Stavroyiannis, S., & Pintelas, P. (2021). An advanced CNN-LSTM model for cryptocurrency forecasting. *Electronics*, 10(3), 287. https://doi.org/10.3390/electronics10030287

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Computing & Applications*, 32(23), 17351–17360. https://doi.org/10.1007/s00521-020-04867-x

Mitchell, T. (1997). *Machine Learning.* New York: McGraw Hill.

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications*, 42(4), 2162-2172. https://doi.org/10.1016/j.eswa.2014.10.031

Persio, L. D., & Honchar, O. (2016). Artificial neural networks architectures for stock price prediction: Comparisons and applications. *International Journal of Circuits, Systems and Signal Processing*, 10, 403-413. https://doi.org/10.1007/s40745-020-00305-w

Poorzaker Arabani, S., & Ebrahimpour Komleh, H. (2019). The optimization of forecasting ATMs cash demand of Iran banking network using LSTM deep recursive neural network. *Journal of Operational Research and Its Applications*, 16(3), 69-88. (In Persian)

Raei, R., & Chavoshi, K. (2003). Predicting the stock price: Artificial neural network and multi-factor model. *Financial Research Journal*, 15(5), 97-120. (In Persian)

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing Journal*, 90, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Sezer, O. B., & Ozbayoglu, A. M. (2018). Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. *Applied Soft Computing Journal*, 70, 525-538. https://doi.org/10.1016/j.asoc.2018.04.024

Shoushtarian, B., Zahedi, E., Yazdinejad, M., & Hormozan, S. (2018). Predicting the price of precious metal by deep learning algorithms. *International Conference on Innovation & Research on Engineering Sciences. Tbilisi.* (In Persian)

Zhou, F., Zhou, H. M., Yang, Z., & Yang, L. (2019). EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction. *Expert Systems with Applications*, 115, 136-151. https://doi.org/10.1016/j.eswa.2018.07.065

Zhou, X., Pan, Z., Hu, G., Tang, S., & Zhao, C. (2018). Stock market prediction on high frequency data using generative adversarial nets. *Mathematical Problems in Engineering*, 2018, 1–11.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مدیریت دارایی و تأمین مالی
سال دهم، شماره سوم، شماره پیاپی (۳۸)، پاییز ۱۴۰۱، ص ۱-۲۰
تاریخ دریافت: ۱۴۰۰/۰۴/۰۷ تاریخ پذیرش: ۱۴۰۰/۱۲/۱۶

مقاله پژوهشی

کاربرد معماری‌های یادگیری عمیق در پیش‌بینی قیمت سهام (رویکرد شبکهٔ عصبی پیچشی CNN)

امیر شریف فر
دانشجوی دکتری، گروه مدیریت مالی، دانشکدهٔ مدیریت و اقتصاد، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
am.sharifar@gmail.com
مریم خلیلی عراقی *
استادیار، گروه مدیریت بازرگانی، دانشکدهٔ مدیریت و اقتصاد، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
m.khaliliaraghi@srbiau.ac.ir
ایمان رئیسی وانانی
دانشیار، گروه مدیریت صنعتی، دانشکدهٔ مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران
imanraeesi@atu.ac.ir
میرفیض فلاح شمس
دانشیار، گروه مدیریت مالی، دانشکدهٔ مدیریت، واحد تهران مرکز، دانشگاه آزاد اسلامی، تهران، ایران
fallahshams@gmail.com

چکیده

اهداف: الگوریتم‌های مبتنی بر شبکهٔ عصبی پیچشی (CNN) که شاخه‌ای از مبحث یادگیری عمیق است، در سال‌های اخیر پیشرفت چشمگیری در حوزه‌های تحلیل فیلم و تصویر داشته‌اند؛ موفقیت و پذیرفته‌شدن الگوهای نوین این حوزه باعث به‌کارگیری گستردهٔ آنها در زمینه‌های مختلف اعم از تحلیل متن و داده‌های سری زمانی شده است. یادگیری عمیق بخشی از الگوریتم‌های یادگیری ماشینی است که در آن از چندین لایهٔ پردازش اطلاعات به‌ویژه اطلاعات غیرخطی استفاده می‌شود تا از ورودی خام، بهترین ویژگی‌های مناسب با هدف تحلیل، بازشناخت الگو یا پیش‌بینی استخراج شود.

روش: در پژوهش حاضر توانایی معماری‌های مختلف الگوریتم CNN برای پیش‌بینی قیمت سهام بررسی شده است.

نتایج: نتایج حاصل از اجرای الگوریتم به تعداد ۵۴ دفعه با معماری‌ها و پارامترهای متفاوت و با استفاده از دو دستهٔ اصلی داده‌های ورودی شامل اطلاعات قیمتی روزانهٔ سهام و ده شاخص منتخب تکنیکال برای سهام شرکت ذوب‌آهن اصفهان نشان‌دهندهٔ آن است که استفاده از CNN همراه با لایهٔ ادغام بیشینه (ترکیب پارامترهای اندازهٔ دستهٔ ۶۴، تعداد فیلتر ۲۵۶ و با تابع فعال‌سازی ReLU)، دارای خطاهای درصد MAPE = ۱/۷۹ و درصد NRMSE = ۲/۷۱ است که نشان‌دهندهٔ عملکرد بهتر آن نسبت‌به سایر معماری‌ها و الگوریتم RNN است.

کلیدواژه‌ها: پیش‌بینی قیمت سهام، یادگیری ماشینی، یادگیری عمیق، الگوریتم شبکهٔ عصبی پیچشی (CNN)، الگوریتم شبکهٔ عصبی بازگشتی (RNN).

---

* نویسندهٔ مسئول

شریف فر، امیر.، خلیلی عراقی، مریم.، رئیسی وانانی، ایمان. و فلاح شمس، میرفیض. (۱۴۰۱). کاربرد معماری‌های یادگیری عمیق در پیش‌بینی قیمت سهام (رویکرد شبکهٔ عصبی پیچشی CNN). *مدیریت دارایی و تأمین مالی*، ۱۰ (۳)، ۱-۲۰.

2383-1189 / © 2022

This is an open access article Under the BY-NC-ND/4.0/ License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

http://dx.doi.org/10.22108/AMF.2022.129205.1673 https://dorl.net/dor/20.1001.1.23831189.1401.10.3.1.9

مقدمه

فعالین بازار سرمایه به‌منظور کسب بالاترین بازده از سرمایه‌گذاری خود همواره به‌دنبال پیش‌بینی قیمت سهام‌اند و این امر توجه بسیاری از پژوهش‌گران دانشگاهی و کارشناسان این حوزه را به خود جلب کرده است؛ به‌طور کلی فرآیندهای سری زمانی را می‌توان به دو طبقهٔ خطی و غیرخطی دسته‌بندی کرد که بر این اساس الگوهای خطی با روش‌های ساده و الگوهای غیرخطی (تصادفی و آشوبناک) با الگوریتم‌های پیچیده و قوی قابل پیش‌بینی است؛ الگوریتم‌های قیمت‌گذاری دارایی‌های سرمایه‌ای (CAPM)[1]، سه عاملی فاما و فرنچ و قیمت‌گذاری آربیتراژ (APT)[2] از نظریه‌های بنیادی و اولیه حوزهٔ مالی است که با فرض خطی‌بودن رابطهٔ بین متغیرها سعی در پیش‌بینی بازده دارد (Gordon & Francis, 1986)؛ اما از آنجا که مهم‌ترین ویژگی سری‌های زمانی مالی مانند: قیمت سهام، نرخ تورم، نرخ ارز و قیمت نفت غیرخطی‌بودن آنهاست، استفاده از روش‌های خطی برای پیش‌بینی سری‌های زمانی مالی نتایج مطلوبی به همراه نخواهد داشت؛ از این رو، درخصوص مسائلی همچون پیش‌بینی قیمت در بازار سرمایه، اغلب الگوریتم‌های مبتنی بر رویکرد بی‌نظمی و پویایی غیرخطی نظیر یادگیری ماشین[3] ازجمله روش‌های مورد اقبال پژوهش‌گران در سال‌های اخیر بوده است (Raei, 2003).

یادگیری ماشین که زیرمجموعهٔ هوش مصنوعی است، به علت خاصیت اکتشافی خود، بدون هیچ فرض اولیه‌ای شروع به الگوسازی رفتار داده‌ها می‌کند و به‌مرور زمان و با جلو رفتن الگوریتم، الگو پررنگ‌تر خواهد شد؛ ساختار غیرخطی و مقاوم این روش، توانایی شبیه سازی رفتار محیط واقعی را فراهم کرده است (Mitchell, 1997). در یادگیری ماشینی از طریق به‌کارگیری یک سری از الگوریتم‌های خاص سعی می‌شود الگوهای پنهان بین داده‌ها کشف شود؛ در صورتی که تعداد متغیرهای مستقل و وابسته زیاد باشد و بین آنها ارتباط خطی برقرار نباشد، این روش در زمرهٔ بهترین گزینه‌ها برای پیش‌بینی است.

یادگیری عمیق، زیرشاخه‌ای از یادگیری ماشینی و از حوزه‌های نوین این بخش است. یادگیری عمیق معرف سلسله‌ای از الگوریتم‌هاست که در آن از چندین لایهٔ پردازش اطلاعات به‌ویژه اطلاعات غیرخطی استفاده می‌شود تا بهترین ویژگی‌های مناسب از ورودی خام استخراج شود (Deng & Yu, 2014). از میان تکنیک‌های مختلف یادگیری عمیق که در علوم مختلف کاربردهای فراوانی دارد، پژوهش‌گران حوزهٔ مالی برای پیش‌بینی قیمت با توجه به ویژگی‌های سری‌های زمانی مالی، اغلب از الگوریتم‌های به‌خصوصی نظیر شبکهٔ عصبی بازگشتی (RNN)[4]، حافظهٔ کوتاه‌مدت ماندگار (LSTM)[5] و شبکهٔ عصبی پیچشی (CNN)[6] استفاده کرده‌اند (Sezer and et al., 2020).

در این پژوهش ضمن معرفی انواع روش‌های یادگیری عمیق که در پیش‌بینی سری‌های زمانی مالی استفاده می‌شود، مقایسهٔ توانایی معماری‌های مختلف الگوریتم CNN در پیش‌بینی قیمت سهام بررسی شده است. طبق بررسی‌های انجام‌شده الگوریتم CNN معماری‌های متفاوتی دارد و در هریک از پژوهش‌های انجام‌شده در گذشته اغلب بر یکی از معماری‌ها تمرکز شده است؛ در حالی که در این مطالعه در سطح اول معماری‌های مختلف CNN استفاده‌شده در پژوهش‌های مختلف استخراج شده و در سطح دوم هریک از معماری‌های شناسایی شده با پارامترهای مختلف بهینه سازی شده است. درنهایت بهترین عملکرد هر معماری با پارامترهای مختلف برای انتخاب معماری بهینه مقایسه شده است که در مطالعه‌های گذشته این موضوع با جامعیت حاضر بررسی نشده است.

---

1. Capital Asset Pricing Model
2. Arbitrage Pricing Theory
3. Machine Learning
4. Recurrent Neural Network
5. Long Short-Term Memory
6. Convolutional Neural Network

مبانی نظری

به موازات کاربرد انواع روش‌های مختلف برای پیش‌بینی سری‌های زمانی مالی، پژوهش‌گران حوزهٔ یادگیری ماشینی نیز با استفاده از الگوریتم‌های متنوع موجود در این بخش سعی در پیش‌بینی قیمت داشته‌اند. یادگیری ماشینی به مفهوم مطالعهٔ الگوریتم‌های کامپیوتری است که براساس تجربه می‌آموزد و پیشرفت می‌کند (Mitchell, 1997). این دانش زمینهٔ پژوهشی به‌نسبت جدیدی از هوش مصنوعی است که در گسترهٔ وسیعی از علوم مختلف اعم از کنترل ربات، تشخیص چهره و گفتار، شناسایی متن، پردازش داده‌های اینترنتی، داده‌کاوی و پیش‌بینی سری‌های زمانی کاربرد فراوان دارد. پژوهش‌گران مالی نیز از روش‌های متعددی در حوزهٔ یادگیری ماشینی نظیر محاسبات تکاملی (ECs)[1]، الگوسازی عامل بنیان[2] و شبکهٔ عصبی مصنوعی برای بررسی داده‌های سری زمانی استفاده کرده‌اند (Sezer and et al., 2020).

یادگیری عمیق، گونه‌ای از شبکه‌های عصبی مصنوعی است که شامل چندین لایه پردازش اطلاعات است که به الگوریتم توانایی بیشترین همگامی را با داده‌ها می‌دهد. مهم‌ترین مزیت الگوریتم‌های یادگیری عمیق نسبت‌به روش سنتی شبکهٔ عصبی، استخراج خودکار ویژگی‌های مناسب از ورودی‌های خام است که از آن برای روند یادگیری الگو[3] استفاده می‌شود. مطالعات انجام‌شده در حوزهٔ پیش‌بینی سری‌های زمانی مالی که در آنها از الگوریتم‌های یادگیری عمیق استفاده شده است، براساس نوع خروجی موردانتظار اغلب به دو دستهٔ اصلی تقسیم می‌شود که یکی پژوهش‌هایی است که با هدف پیش‌بینی قیمت سهام (Chong and et al., 2017; Chen and et al., 2015; Dezsi & Nistor, 2016) و دیگری با هدف پیش‌بینی روند قیمت (صعودی، نزولی یا خنثی) انجام شده است (Das and et al., 2018;. Zhou and et al., 2019).

در پژوهش‌های مرتبط با هدف پیش‌بینی قیمت سهام، از روش‌های متنوعی در حوزهٔ یادگیری عمیق استفاده شده است که از آن جمله می‌توان به الگوریتم شبکهٔ عصبی مکرر یا بازگشتی (RNN) (and et al., 2018; Lee & Yoo, 2020)، حافظهٔ کوتاه‌مدت ماندگار (LSTM) و شبکهٔ عصبی پیچشی (CNN) اشاره کرد که در این میان پژوهش‌گران به الگوریتم‌های LSTM و CNN به دلیل دارا بودن برخی ویژگی‌ها بیشتر از بقیهٔ روش‌ها توجه کرده‌اند؛ برای مثال، گائو[4] و همکاران (2020) مقایسهٔ قدرت انواع الگوریتم‌های LSTM، MLP و CNN را درخصوص پیش‌بینی شاخص‌های S&P500 و Nikkei225 در افق زمانی یک‌روزه بررسی کرده‌اند؛ آنها از متغیرهای ورودی شاخص‌های تکنیکال، داده‌های قیمتی و شاخص‌های کلان اقتصادی برای پیش‌بینی استفاده کرده‌اند. درنهایت با استفاده از خطای RMSE بهترین عملکرد به‌ترتیب الگوریتم‌های CNN، LSTM و MLP بوده است. دیپر سیو و هانچر[5] (2016) نیز با استفاده از الگوریتم‌های CNN، LSTM و MLP، داده‌های تاریخی قیمت پایانی شاخص S&P500 را بررسی کرده‌اند که نتایج حکایت از عملکرد بهتر الگوریتم CNN داشته است. در جدول (۱) برخی از پژوهش‌های انجام‌شده با استفاده از الگوریتم‌های یادگیری عمیق آورده شده است.

---

1. Evolutionary Computations
2. Agent-based models
3. Model
4. Gao
5. Di Persio & Honchar

جدول (۱) مطالعات مرتبط با به‌کارگیری معماری‌های مختلف یادگیری عمیق برای پیش‌بینی قیمت سهام

Table (1) Studies using different deep learning algorithms for stock price forecasting.

حوزه‌های مطالعاتی / پژوهش‌های مرتبط	جامعهٔ آماری	دورهٔ زمانی مورد بررسی	داده‌های مورد استفاده	قالب زمانی	افق زمانی پیش‌بینی	الگوریتم مورد استفاده
ژو[1] و همکاران، (2018)	۴۲ شرکت در SSE	۲۰۱۶	داده‌های قیمتی، شاخص‌های تکنیکال	۲۴۲ دقیقه	یک دقیقه	CNN LSTM,
هیرانشا[2] و همکاران، (2018)	۵ شرکت در NSE	۲۰۱۶–۱۹۹۷	داده‌های قیمتی، گردش مالی، تعداد دفعات معاملات	۲۰۰ روز	یک الی ده روز	LSTM, RNN, CNN, MLP
لی و یو[3] (2020)	۱۰ شرکت در S&P500	۲۰۱۶–۱۹۹۷	داده‌های قیمتی	۳۶ دقیقه	یک دقیقه	RNN, LSTM, GRU
حسین زاده و هراتی‌زاده[4] (2019)	۵ شاخص بازار سهام آمریکا	۲۰۱۷–۲۰۱۰	داده‌های قیمتی، شاخص‌های تکنیکال، قیمت طلا و نفت	–	یک روز	CNN
لیو[5] و همکاران، (2017)	بازار سهام چین	۲۰۱۷–۲۰۰۷	داده‌های قیمتی	۳۰ روز	یک الی پنج روز	CNN, LSTM
خارو[6] و همکاران، (2017)	۱۰ شرکت در NYSE	۲۰۱۷	داده‌های قیمتی، شاخص‌های تکنیکال	بیست دقیقه	یک دقیقه	LSTM, MLP
گوندوز[7] و همکاران، (2017)	۱۰۰ شرکت در بازار سهام استانبول	۲۰۱۵–۲۰۱۱	داده‌های قیمتی، شاخص‌های تکنیکال	۱۰ روز	یک روز	CNN

در بخش داخلی، مطالعات انگشت شماری در حوزهٔ کاربرد معماری‌های مختلف یادگیری عمیق در مباحث مالی انجام شده است. ازجمله پژوهش شوشتریان[8] و همکاران (۱۳۹۷) که سعی در پیش‌بینی قیمت فلزات گران‌بها با استفاده از دو روش LSTM و RNN داشتند؛ نتایج پژوهش حاکی از برتری روش RNN است. علاوه بر این، دامی و خدابخش[9] (۱۳۹۷) با استفاده از بازنمایی‌های عددی و متنی با روش LSTM اثرات زمانی وقایع گذشته را بر قیمت‌های موجود بررسی کرده‌اند.

روش شبکهٔ عصبی پیچشی (CNN) یک شبکهٔ عصبی مصنوعی (ANN) پیش‌خور محسوب می‌شود که ورودی‌های آن به‌صورت ماتریس دوبعدی است و برخلاف یک شبکهٔ عصبی دارای اتصال کامل مانند MLP، موقعیت داده‌ها در بردار ورودی (یا ماتریس) حائز اهمیت است (Sezer & Ozbayoglu, 2018). به دلیل ویژگی تحلیل ماتریسی شبکهٔ کانولوشنی یا پیچشی،[10] این روش کاربرد وسیعی در حوزه‌های گوناگون اعم از تشخیص تصویر، فیلم، پردازش زبان طبیعی و سیستم‌های خبره دارد. یکی از زمینه‌هایی که به‌تازگی موردتوجه پژوهش‌گران قرار گرفته است، تحلیل سری‌های زمانی مانند قیمت سهام با استفاده از این روش نوین است. در معماری CNN از لایه‌های مختلفی استفاده می‌شود که به‌ترتیب عبارت است از: پیچشی، ادغام،[11] حذف تصادفی[12] و تمام متصل.[13] در رابطه با عملکرد CNN به‌طور خلاصه می‌توان گفت در سطح اول، لایهٔ پیچشی قرار دارد که ورودی آن یک آرایه از اعداد است؛ در این لایه با

1. Zhou
2. Hiransha
3. Lee & Yoo
4. Hoseinzade & Haratizadeh
5. Liu
6. Khare
7. Gunduz
8. Shoushtarian
9. Dami & Khodabakhsh
10. Convolution
11. Pooling
12. Dropout
13. Fully Connect or Dense Layer

استفاده از فیلترهای[1] متنوع ویژگی‌های جدیدی از ماتریس استخراج می‌شود.[2] شایان ذکر است فیلترها نیز خود آرایه‌ای از وزن‌ها هستند که در حین مراحل آموزش به‌روز می‌شوند و به‌دنبال آن عملیات ادغام انجام می‌شود که در آن وظیفهٔ کاهش ابعاد و تعداد پارامترهای شبکه را انجام می‌شود (Sezer and et al., 2020). شبکه‌های عصبی عمیق اغلب در معرض بیش‌برازش[3] هستند؛ چون تعداد لایه‌های اضافه شده این امکان را ایجاد می‌کند که وابستگی‌های نایاب در داده‌های آموزشی حساب شود. با استفاده از روش حذف تصادفی می‌توان به از بین بردن بیش‌برازش کمک کرد؛ در این شیوه درصد مشخصی از واحدها به‌صورت تصادفی از لایه‌های مخفی حذف می‌شود. این کار به از بین بردن وابستگی‌های نادر در داده‌های آموزشی کمک می‌کند (Poorzaker Arabani & Ebrahimpour Komleh, 2019). خروجی این لایه بعد از تبدیل به بردار یک‌بعدی به لایهٔ شبکه اتصال کامل ارسال و در آن از الگوریتم‌های رایج شبکه‌های عصبی استفاده می‌شود. عملکرد لایهٔ پیچشی در فرمول زیر نشان داده شده است.

$$s(t) = (x * w)(t) = \sum_{i=-\infty}^{\infty} x(i)w(t-i) \qquad (۱)$$

در رابطهٔ (۱) متغیرهای t، s، w، x و i به‌ترتیب نشان‌دهندهٔ زمان، نقشهٔ ویژگی[4] فیلتر، داده ورودی و شمارنده هستند. در این پژوهش با تمرکز بر معماری‌های مختلف الگوریتم CNN به پرسش‌های زیر پاسخ داده می‌شود. پیش‌بینی قیمت سهام در بورس اوراق بهادار تهران با استفاده از معماری‌های مختلف الگوریتم CNN چگونه است؟ کدام معماری CNN عملکرد بهتری در پیش‌بینی قیمت سهام دارد؟ درنهایت اینکه عملکرد معماری‌های منتخب الگوریتم CNN در مقایسه با الگوریتم RNN چگونه است؟

روش پژوهش

سهام شرکت ذوب‌آهن اصفهان به دلیل حجم معاملات روزانهٔ به‌نسبت بالا در مقایسه با سایر شرکت‌ها و همچنین درصد مناسب سهام شناور آزاد برای تحلیل انتخاب شده است. از این رو، نمونهٔ آماری، قیمت‌های روزانهٔ شرکت ذوب‌آهن اصفهان از سال ۱۳۹۶ تا سال ۱۳۹۹ است. در راستای لزوم هم‌سان‌بودن دورهٔ زمانی متغیرها و افق زمانی پیش‌بینی قیمت در پژوهش (که به صورت روزانه است)، این مهم در انتخاب متغیرهای ورودی در نظر گرفته شده است و از دو دسته داده‌های قیمتی (اولین قیمت، بالاترین قیمت، پایین‌ترین قیمت، آخرین قیمت و حجم معاملات) و شاخص‌های تکنیکال استفاده شده است که مجموع متغیرهای داده‌های قیمتی از نرم‌افزار TseClient و شاخص‌های تکنیکال براساس پژوهش‌های کارا[5] و همکاران (2011) و پاتل[6] و همکاران (2015) به شرح جدول (۲) محاسبه شده است. درنهایت برای طراحی الگو از نرم‌افزار Python نسخهٔ ۳.۸ همراه با توابع کتابخانه‌ای کراس استفاده شده است.

1. Filter or Kernel
2. Feature Extraction
3. Over Fitting
4. Feature map
5. Kara
6. Patel

جدول (۲) معادلات شاخص‌های تکنیکال

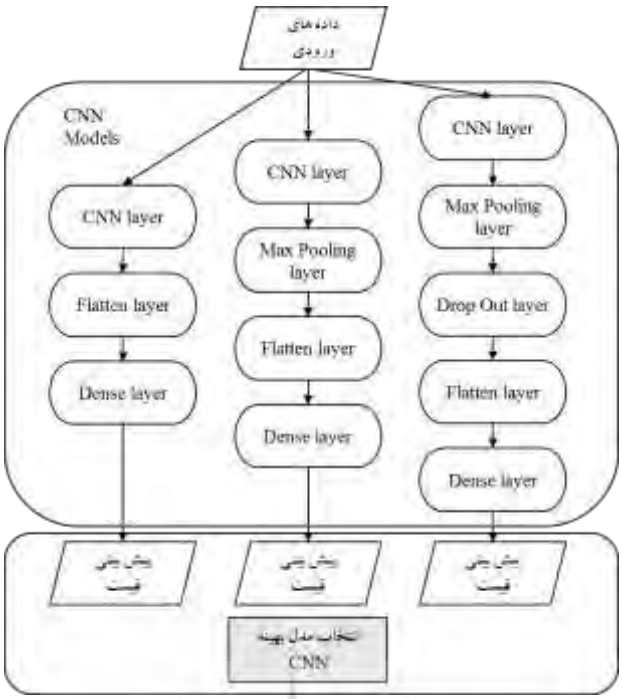
Table (2) Technical indicators equation.

نام شاخص	نحوهٔ محاسبه
Weighted 10-day moving average	$\frac{(10)C_t + (9)C_{t-1} + \cdots + C_{t-9}}{n + (n-1) + \cdots + 1}$
Momentum	$C_t - C_{t-9}$
Stochastic K%	$\frac{C_t - LL_{t-(n-1)}}{HH_{t-(n-1)} - LL_{t-(n-1)}} \times 100$
Stochastic D%	$\frac{\sum_{i=0}^{n-1} K_{t-i}}{10}\%$
Relative strength index (RSI)	$100 - \frac{100}{1 + (\frac{\sum_{i=0}^{n-1} UP_{t-i}}{n})/(\frac{\sum_{i=0}^{n-1} DW_{t-i}}{n})}$
Moving average convergence divergence (MACD)	$MACD(n)_{t-1} + \frac{2}{n+1} \times (DIFF_t - MACD(n)_{t-1})$
MACD Signal	$\left(MACD * \frac{2}{n+1}\right) + (Signal_{perv} * \left(1 - \frac{2}{n+1}\right))$
Larry William's R%	$\frac{H_n - C_t}{H_n - L_t} \times 100$
On Balance Volume (OBV)	$OBV_{prev} + \begin{cases} volume. & if\ close > close_{prev} \\ 0. & if\ close = close_{prev} \\ -volume. & if\ close < close_{prev} \end{cases}$
CCI (Commodity channel index)	$\frac{M_t - SM_t}{0.015D_t}$

که در آن Ct آخرین قیمت در زمان t، Ht بالاترین قیمت در زمان t، UPt تغییرات قیمت صعودی در زمان t، Lt پایین‌ترین قیمت در زمان t، HHt , LLt بیشترین و کمترین قیمت در t روز آخر و DWt تغییرات قیمت نزولی در زمان t است. علاوه بر این، نحوهٔ محاسبهٔ Mt ، SMt و Dt به شرح ذیل است.

$$M_t = \frac{H_t + L_t + C_t}{3} \qquad SM_t = \frac{(\sum_{i=1}^{n} M_{t-i+1})}{n} \qquad D_t = \frac{(\sum_{i=1}^{n} |M_{t-i+1} - SM_t|)}{n}$$

بررسی پژوهش‌های مرتبط با پیش‌بینی قیمت با استفاده از روش CNN نشان‌دهندهٔ آن است که به‌طور معمول از سه معماری (لایه‌چینی) به شرح شکل (۱) استفاده شده است که در این پژوهش نیز هر سه معماری بررسی شده است.



شکل (۱) فرآیند انتخاب معماری بهینه CNN

Figure (1) The process of choosing optimal CNN algorithm

شـایان ذکر اسـت انتخاب معماری‌های مختلف و پارامترهای مورد اسـتفاده با مدنظر قراردادن مطالعات مختلف انجام‌شـده در این حوزه به شرح جدول ذیل است:

جدول (۳) مطالعات مؤثر در طراحی الگوی پژوهش

Table (3) Effective studies in model development

پژوهش	معماری مورد استفاده	حوزهٔ مطالعاتی	داده‌های مورد استفاده	افق زمانی پیش‌بینی
لوریس[1] و همکاران، (2020)	استفاده از دو لایهٔ پیچشی با تعداد فیلترهای متفاوت	پیش‌بینی دلاری قیمت طلا	داده‌های قیمتی	یک روز
گائو[2] و همکاران، (2020)	CNN ساده CNN همراه با لایهٔ حذف تصادفی	پیش‌بینی شاخص S&P500 CSI300 Nikkei225	داده‌های قیمتی، حجم معاملات، شاخص‌های تکنیکال	یک روز
گودلک[3] و همکاران، (2017)	CNN همراه با لایه‌های حذف تصادفی و ادغام بیشینه[4]	پیش‌بینی قیمت صندوق‌های سرمایه‌گذاری ETF	داده‌های قیمتی، شاخص‌های تکنیکال	یک روز
جی[5] و همکاران، (2019)	CNN همراه با لایهٔ ادغام بیشینه	پیش‌بینی قیمت قردادهای آتی کربن	داده‌های قیمتی	هفت روز
لی و دای[6] (۲۰۲۰)	CNN همراه با لایهٔ ادغام بیشینه	پیش‌بینی قیمت رمزارز بیت‌کوین	داده‌های قیمتی	یک روز

1. Livieris
2. Gao
3. Gudelek
4. Max-Pooling
5. Ji
6. Li & Dai

برای انتخاب معماری بهینهٔ CNN نیاز است که هر سه روش استفاده شده با پارامترهای مختلف بررسی شود. شایان ذکر است پارامترهای اثرگذار در ساخت معماری بهینهٔ CNN شامل مواردی نظیر تعداد فیلترها[1]در لایهٔ CNN، اندازهٔ دسته[2]و تابع فعال‌سازی[3] است. برای مثال، تعداد کم فیلترها باعث ایجاد خطای بیشتری می‌شود و شبکه همگرا نمی‌شود. از طرفی تعداد فیلترهای زیاد به دلیل ایجاد بیش‌برازش منجر به خطای بالا می شود. از این رو، موارد مذکور با توجه به خروجی اخذ شده بهینه سازی شده است. انتخاب پارامترهای مختلف در جدول (۴) براساس نتایج مطالعات انجام‌شده در این حوزه است (Gudelek and et al., 2017; Livieris and et al Pintelas., 2021; Livieris and et al., 2020).

جدول (۴) پارامترهای مختلف مورد استفاده در اجرای الگو

Table (4) The parameters used to run the model

پارامتر	مقادیر
تعداد فیلتر	۶۴، ۱۲۸، ۲۵۶
اندازه دسته	۳۲، ۶۴، ۱۲۸
optimizer	Adam
Loss	Mean Square Error
توابع فعال‌سازی	**رابطهٔ ریاضی**
تابع یکسو ساز (ReLU)	$g(z) = \max(0.z)$
تانژانت هیپربولیک (Tanh)	$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$

برای اجرای الگوریتم، داده‌ها به دو بخش آموزش (۸۰ درصد) برای ساخت الگو و آزمایش (۲۰ درصد) برای آزمون تقسیم شده است؛ برای بهتر شدن فرآیند یادگیری، مرحلهٔ آموزش به تعداد۵۰ مرتبه تکرار[4] شده است. علاوه بر این، باید توجه داشت که تفاوت مقیاس‌ها در متغیرهای مختلف، نتایج را تحت‌تأثیر قرار می‌دهد و به‌طور معمول برای از بین بردن تأثیر نامطلوب مذکور، نرمال‌سازی داده نیاز است که در این پژوهش برای نرمال‌سازی از روش کمینه – بیشینه[5]جهت تبدیل داده‌ها به مجموعهٔ جدیدی استفاده شده است که در آن همهٔ مقادیر بین ۰ و ۱ است.

یافته‌ها

برای انتخاب معماری بهینه CNN، هر سه معماری معرفی شده با پارامترهای مختلف مندرج در جدول (۴) اجرا شده است. درنهایت بهترین نتایج حاصل از هر معماری مقایسه شده است؛ بنابراین معماری منتخب، حاصل از اجرای الگوریتم به تعداد ۵۴ دفعه با لایه‌ها و پارامترهای متفاوت است. در مرحلهٔ اول، روش CNN با معماری لایه‌های ذیل اجرا شده است (معماری اول CNN).

شکل (۲) معماری اول CNN

Figure (2) First CNN architecture

[1]. Filters
[2]. Batch Size
[3]. Activation Function
[4]. Epochs
[5]. Min-Max normalization

در جداول (۵) و (۶) خطاهای میانگین قدر مطلق درصد خطا (MAPE[1]) و مجذور میانگین مربعات خطای نرمال (NRMSE[2]) در اجرای الگو به تعداد ۱۸ دفعه با استفاده از جایگشت‌های مختلف تعداد فیلتر، اندازهٔ دسته و استفاده از دو تابع فعال سازی ReLU و Tanh آورده شده است.

جدول (۵) خطای اجرای معماری اول CNN به تفکیک پارامترهای مختلف با تابع فعال‌سازی ReLU

Table (5) Error of first CNN architecture in terms of different parameters with ReLU function

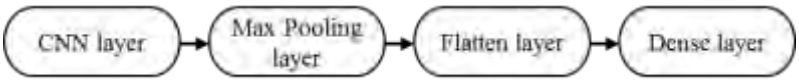
تابع فعال‌سازی ReLU		نوع خطا	تعداد فیلتر		
			۶۴	۱۲۸	۲۵۶
اندازهٔ دسته	۳۲	MAPE	۲/۹۹	۳/۰۰	۱۱/۸۳
		NRMSE	۴/۱۰	۳/۹۴	۱۲/۲۸
	۶۴	MAPE	۴/۷۰	۲/۵۲	۳/۸۵
		NRMSE	۶/۱۸	۳/۵۶	۵/۰۰
	۱۲۸	MAPE	۳/۹۵	۲/۴۳	۳/۱۸
		NRMSE	۴/۹۳	۳/۴۷	۴/۰۹

جدول (۶) خطای اجرای معماری اول CNN به تفکیک پارامترهای مختلف با تابع فعال‌سازی Tanh

Table (6) Error of first CNN architecture in terms of different parameters with Tanh function

تابع فعال‌سازی Tanh		نوع خطا	تعداد فیلتر		
			۶۴	۱۲۸	۲۵۶
اندازهٔ دسته	۳۲	MAPE	۲/۹۹	۲/۵۰	۴/۰۶
		NRMSE	۴/۱۲	۳/۳۳	۴/۳۷
	۶۴	MAPE	۵/۷۴	۲/۳۵	۲/۱۳
		NRMSE	۶/۱۰	۳/۱۹	۳/۰۹
	۱۲۸	MAPE	۴/۰۰	۳/۲۱	۳/۲۵
		NRMSE	۵/۱۵	۴/۳۷	۳/۷۵

نتایج به‌دست‌آمده از اجرای معماری اول حاکی از آن است که ترکیب پارامترها با تعداد فیلتر ۲۵۶، اندازهٔ دسته ۶۴ و با تابع فعال‌سازی Tanh بهترین دقت پیش‌بینی را دارد. در مرحلهٔ دوم، روش CNN با معماری لایه‌های ذیل اجرا شده است (معماری دوم CNN با لحاظ لایهٔ ادغام بیشینه).



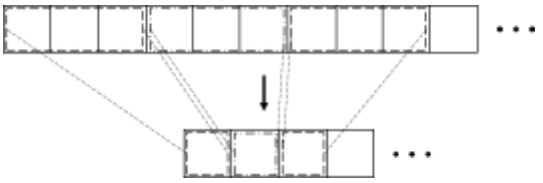
شکل (۳) معماری دوم CNN

Figure (3) Second CNN architecture

قراردادن یک لایهٔ ادغام بین چندین لایهٔ پیچشی پشت سر هم در یک معماری CNN امری رایج است. کارکرد این لایه کاهش تعداد پارامترها و محاسبات در داخل شبکه و کنترل بیش‌برازش[3] است. لایهٔ ادغام به صورت مستقل بر هر برش عمقی از تودهٔ ورودی

[1]. Mean Absolute Percentage Error
[2]. Normalized Root Mean Square Error
[3]. Over Fitting

عمل می‌کند و آن را با استفاده از عملیات بیشینه[1]ازلحاظ مکانی تغییر اندازه[2]می‌دهد. رایج‌ترین نحوهٔ استفاده از این لایه به‌صورت استفادهٔ آن با فیلترهایی با اندازه ۳×۱ به‌همراه گام[3]۳ است که همانند شکل (۴)، عملیات بیشینه در این قسمت، بزرگ‌ترین عدد بین ۳ عدد (یک ناحیه ۳×۱ در برش عمقی) را انتخاب می‌کند و باعث حذف ۶۷ درصد مقادیر موجود در آن برش عمقی می‌شود؛ گفتنی است بُعد عمق بدون تغییر باقی می‌ماند.



شکل (۴) عملکرد لایه ادغام بیشینه

Figure (4) The function of Max-Pooling Layer

نتایج اجرای الگوریتم CNN با لحاظ کردن لایهٔ ادغام بیشینه به شرح جداول (۷) و (۸) است.

جدول (۷) خطای اجرای معماری دوم CNN به تفکیک پارامترهای مختلف با تابع فعال‌سازی ReLU

Table (7) Error of second CNN architecture in terms of different parameters with ReLU function

تابع فعال‌سازی ReLU		نوع خطا	تعداد فیلتر		
			۶۴	۱۲۸	۲۵۶
اندازهٔ دسته	۳۲	MAPE	۲/۷۹	۳/۰۰	۲/۲۶
		NRMSE	۳/۶۵	۴/۳۷	۳/۴۶
	۶۴	MAPE	۲/۴۰	۱/۹۹	۱/۷۹
		NRMSE	۳/۱۶	۲/۸۸	۲/۷۱
	۱۲۸	MAPE	۳/۰۳	۳/۶۶	۲/۱۷
		NRMSE	۴/۳۳	۴/۲۷	۳/۰۵

جدول (۸) خطای اجرای معماری دوم CNN به تفکیک پارامترهای مختلف با تابع فعال‌سازی Tanh

Table (8) Error of second CNN architecture in terms of different parameters with Tanh function

تابع فعال‌سازی Tanh		نوع خطا	تعداد فیلتر		
			۶۴	۱۲۸	۲۵۶
اندازه دسته	۳۲	MAPE	۲/۸۹	۲/۶۱	۵/۲۵
		NRMSE	۴/۱۰	۳/۲۲	۷/۴۶
	۶۴	MAPE	۴/۰۱	۲/۸۱	۷/۲۷
		NRMSE	۴/۹۹	۳/۴۷	۹/۷۸
	۱۲۸	MAPE	۳/۶۴	۳/۰۸	۳/۸۵
		NRMSE	۴/۵۲	۳/۹۱	۴/۳۹

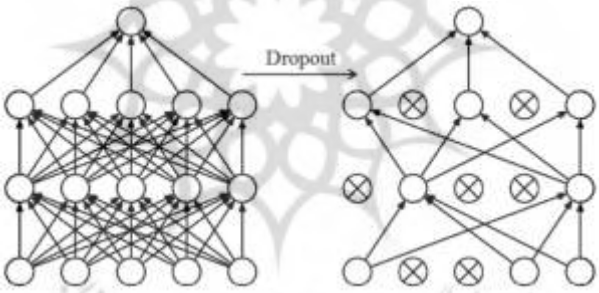
[1]. Max
[2]. Resize
[3]. Stride

نتایج به‌دست‌آمده از اجرای معماری دوم نشان‌دهندهٔ آن است که ترکیب پارامترها با اندازهٔ دسته ۶۴، تعداد فیلتر ۲۵۶ و با تابع فعال‌سازی ReLU بهترین دقت پیش‌بینی را دارد؛ همان‌گونه که مشخص است خطای به‌دست‌آمده با پارامترهای مذکور از نتایج منتخب معماری اول نیز به‌مراتب بهتر است. در مرحلهٔ سوم، روش CNN با معماری لایه‌های ذیل اجرا شده است (معماری سوم CNN با لحاظ لایهٔ ادغام بیشینه و حذف تصادفی).

شکل (۵) معماری سوم CNN

Figure (5) Third CNN architecture

از معایب مربوط به شبکه‌های عصبی عمیق، مشکل بیش‌برازش و زمان محاسباتی بالا ست. شبکه‌های عصبی عمیق در معرض بیش‌برازش هستند؛ چون تعداد لایه‌های اضافه شده این امکان را ایجاد می‌کند که وابستگی‌های نایاب در داده‌های آموزشی حساب شود. با استفاده از روش حذف تصادفی می‌توان به از بین بردن بیش‌برازش کمک کرد؛ در این شیوه درصد مشخصی از واحدها به‌صورت تصادفی از لایه‌های مخفی حذف می‌شود. این کار به از بین بردن وابستگی‌های نادر در داده‌های آموزشی کمک می‌کند (Poorzaker and et al., 2019). در شکل (۶) نحوهٔ عملکرد لایهٔ حذف تصادفی نشان داده می‌شود.



شکل (۶) عملکرد لایه حذف تصادفی (Cao and et al., 2019)

Figure (6) The function of Dropout Layer (Cao and et al., 2019)

نتایج اجرای الگوریتم CNN با لحاظ‌کردن لایهٔ حذف تصادفی و ادغام بیشینه به شرح جداول (۹) و (۱۰) است.

جدول (۹) خطای اجرای معماری سوم CNN به تفکیک پارامترهای مختلف با تابع فعال‌سازی ReLU

Table (9) Error of third CNN architecture in terms of different parameters with ReLU function

تابع فعال‌سازی ReLU		نوع خطا	تعداد فیلتر		
			۶۴	۱۲۸	۲۵۶
اندازهٔ دسته	۳۲	MAPE	۴/۵۷	۲/۷۲	۴/۵۷
		NRMSE	۵/۶۸	۴/۱۲	۵/۴۹
	۶۴	MAPE	۳/۷۶	۳/۲۴	۲/۱۸
		NRMSE	۴/۸۴	۴/۲۹	۳/۲۶
	۱۲۸	MAPE	۲/۹۷	۳/۴۹	۴/۵۲
		NRMSE	۴/۰۲	۴/۴۶	۵/۴۶

جدول (۱۰) خطای اجرای معماری سوم CNN به تفکیک پارامترهای مختلف با تابع فعال‌سازی Tanh

Table (10) Error of third CNN architecture in terms of different parameters with Tanh function

تابع فعال‌سازی Tanh		نوع خطا	تعداد فیلتر ۶۴	تعداد فیلتر ۱۲۸	تعداد فیلتر ۲۵۶
اندازهٔ دسته	۳۲	MAPE	۳/۸۸٪	۷/۷۷٪	۳/۸۲٪
		NRMSE	۵/۰۰٪	۹/۴۸٪	۵/۴۰٪
	۶۴	MAPE	۸/۱۷٪	۲/۷۱٪	۳/۰۱٪
		NRMSE	۸/۹۴٪	۳/۴۳٪	۴/۲۶٪
	۱۲۸	MAPE	۳/۴۴٪	۳/۶۰٪	۳/۷۷٪
		NRMSE	۴/۳۲٪	۴/۸۰٪	۴/۶۳٪

نتایج به‌دست‌آمده از اجرای معماری سوم حاکی از آن است که ترکیب پارامترها با اندازهٔ دسته ۶۴، تعداد فیلتر ۲۵۶ و با تابع فعال‌سازی ReLU بهترین دقت پیش‌بینی را در این بخش دارد.

همان‌طور که در بخش‌های قبلی اشاره شد، RNN یکی دیگر از الگوریتم‌های یادگیری عمیق است که استفاده از آن در پیش‌بینی سری‌های زمانی امری متداول است؛ از این رو، در جدول (۱۱) نتایج حاصل از معماری‌های منتخب و الگوریتم RNN آورده شده است. برای ارزیابی عملکرد، معیارهای MAPE و NRMSE معماری‌های منتخب اول، دوم و سوم CNN با الگوریتم RNN مقایسه شده است.

جدول (۱۱) مقایسهٔ خطای اجرای الگوهای منتخب

Table (11) Comparing the error of selected models

نام روش / نوع خطا	MAPE	NRMSE
RNN	۲/۴۶	۲/۷۹
منتخب معماری اول CNN	۲/۱۳	۳/۰۹
منتخب معماری دوم CNN	۱/۷۹	۲/۷۱
منتخب معماری سوم CNN	۲/۱۸	۳/۲۶

براساس نتایج جدول (۱۱)، دقت منتخب معماری دوم CNN نسبت‌به روش‌های دیگر بالاتر است. مقادیر NRMSE برای هر چهار روش کمتر از سطح ۰/۰۵ است که نشان از برازش مناسب الگوریتم‌های مذکور دارد (Kiyani Mavi & Sayadi Nik, 2015).

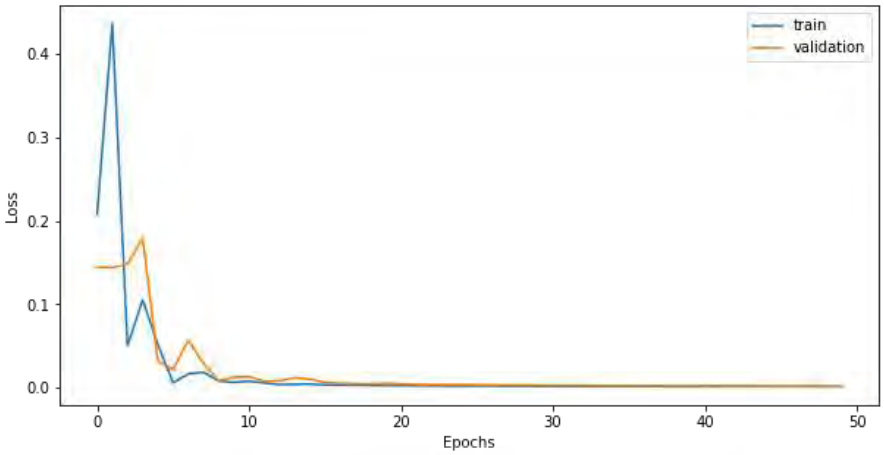
در ادامه به‌منظور بررسی رابطهٔ بین مقدار واقعی قیمت سهام و مقدار پیش‌بینی‌شده از ضریب همبستگی پیرسون استفاده شده است که نتایج آن در جدول (۱۲) آورده شده است. با توجه به اینکه سطح معناداری برای هر چهار روش از سطح خطای یک درصد کمتر است، وجود همبستگی بین مقدار واقعی قیمت سهام و مقدار پیش‌بینی‌شده به لحاظ آماری تأیید شده است.

جدول (۱۲) ضریب همبستگی پیرسون بین مقدار واقعی قیمت سهام و مقدار پیش‌بینی‌شده

Table (12) Pearson's correlation coefficient between real & predicted stock price

شرح / نام روش	RNN	منتخب معماری اول CNN	منتخب معماری دوم CNN	منتخب معماری سوم CNN
ضریب همبستگی پیرسون	۰/۹۲۲	۰/۹۸۲	۰/۹۸۵	۰/۹۰۳
سطح معناداری	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰
تعداد دادهٔ آزمایش	۹۱	۹۱	۹۱	۹۱

نظر به اینکه براساس نتایج جدول (۱۱)، منتخب معماری دوم CNN نسبت‌به روش‌های دیگر عملکرد بهتری داشته است، در ادامه نتایج حاصل از اجرای این روش آورده شده است. فرآیند کاهش خطا در مرحلهٔ آموزش معماری منتخب با تکرار ۵۰ مرتبه بهینه‌سازی در شکل (۷) مشخص است.



شکل (۷) فرآیند کاهش خطا برای مرحلهٔ آموزش در معماری منتخب CNN

Figure (7) The process of error reduction for selected CNN architecture in train step

نمودار قیمت واقعی سهام شرکت ذوب‌آهن ا صفهان به‌همراه قیمت پیش‌بینی به تفکیک داده‌های آموزش و آزمایش در شکل (۸) نشان داده شده است:

شکل (۸) داده‌های واقعی قیمت سهام شرکت ذوب‌آهن اصفهان به‌همراه نمودار پیش‌بینی با استفاده از معماری منتخب

Figure (8) The chart of real stock price compared with the prediction graph using selected model (ZobAhan Esfahan stock price)

نتایج و پیشنهادها

به موازات انواع روش‌های مختلف استفاده‌شده برای پیش‌بینی قیمت سهام، پژوهش‌گران حوزهٔ یادگیری ماشینی نیز با استفاده از الگوریتم‌های متنوع موجود در این بخش سعی در پیش‌بینی قیمت دا شته‌اند. یادگیری عمیق، گونه‌ای از الگوریتم‌های یادگیری ما شینی است که تلاش می‌کند مفاهیم انتزاعی سطح بالا را با استفاده از یادگیری در سطوح و لایه‌های مختلف لحاظ کند و شامل چندین لایه پردازش اطلاعات است که باعث بیشترین همگامی الگوریتم با داده‌ها می‌شود. در پژوهش حاضر از سه معماری مختلف CNN با

ترکیب پارامترهای مختلف برای انتخاب الگوریتم بهینه پیش‌بینی قیمت استفاده شده است؛ بررسی‌ها حاکی از آن بود که پارامترهای اثرگذار در ساخت معماری بهینه CNN شامل مواردی نظیر اندازهٔ دسته، تعداد فیلتر و تابع فعال‌سازی است؛ از این رو، در مرحلهٔ اول با استفاده از ترکیبات مختلف پارامترهای مذکور، مقادیر بهینه برای هر معماری به دست آمد و در مرحلهٔ نهایی مقایسهٔ نتایج حاصل از هر سه معماری منتخب نشان‌دهندهٔ آن بود که منتخب معماری دوم CNN همراه با لایهٔ ادغام بیشینه با ترکیب پارامترهای اندازهٔ دسته ۶۴، تعداد فیلتر ۲۵۶ و با تابع فعال‌سازی ReLU، دارای خطاهای درصد MAPE = ۱/۷۹ و درصد NRMSE = ۲/۷۱ است که نشان‌دهندهٔ عملکرد بهتر آن نسبت‌به سایر معماری‌ها و روش RNN است. معماری به‌دست‌آمده از این پژوهش با نتایج حاصل از مطالعات جی و همکاران (2019) برای پیش‌بینی قیمت قراردادهای آتی کربن و همکاران (2020) درخصوص پیش‌بینی قیمت رمزارزها، در استفاده از لایهٔ پیچشی همراه با لایهٔ ادغام بیشینه مطابقت دارد. بنابراین در مطالعهٔ گائو و همکاران (2020) از لایهٔ پیچشی همراه با لایهٔ حذف تصادفی استفاده شده است و گودلک و همکاران (2017) نیز برای پیش‌بینی قیمت صندوق‌های سرمایه‌گذاری ETF پیشنهاد استفاده از لایهٔ پیچشی را همراه با لایه‌های حذف تصادفی و ادغام بیشینه دادند که در این پژوهش با بررسی معماری‌های بالا، نتایج آنها تأیید نشد و دقت معماری منتخب در استفاده از لایهٔ پیچشی همراه با لایهٔ ادغام بیشینه بهتر از معماری‌های دیگر بوده است.

در پژوهش حاضر با توجه به افق زمانی پیش‌بینی که یک روز آتی است، از دو دسته متغیرهای ورودی شامل داده‌های قیمتی (اولین قیمت، بالاترین قیمت، پایین‌ترین قیمت، آخرین قیمت و حجم معاملات) و ده شاخص تکنیکال استفاده شده است. حال آنکه در پژوهش‌های آتی می‌توان از طیف گسترده‌تری از متغیرهای ورودی نظیر متغیرهای بنیادی استفاده کرد و تأثیر هر گروه از انواع مختلف ورودی‌ها را بر دقت پیش‌بینی بررسی کرد یا به پیش‌بینی قیمت در افق‌های زمانی مختلف پرداخت و توانایی الگو را در پیش‌بینی افق‌های زمانی مختلف با یکدیگر مقایسه کرد. علاوه بر این، می‌توان عملکرد روش پیشنهادی این پژوهش را با سایر الگوریتم‌های هوش مصنوعی مقایسه کرد. از طرفی، تعداد بالای متغیرهای ورودی به‌طور معمول باعث افزایش پیچیدگی و به تبع آن کاهش کارایی الگو می‌شود. از این رو، استفاده از تکنیک‌های کاهش ابعاد داده‌های ورودی ممکن است به افزایش دقت الگوریتم کمک کند. در پژوهش‌های آتی می‌توان از تکنیک‌های مختلف کاهش بعد مانند تجزیه‌وتحلیل مؤلفه‌های اساسی (PCA)[1]، تجزیهٔ مقادیر منفرد(SVD)[2] یا تحلیل تفکیک‌کنندهٔ تعمیم یافته (GDA)[3]استفاده و نتایج را در افزایش دقت الگوریتم بررسی کرد. برخی از پژوهش‌گران نیز با ترکیب الگوریتم CNN با سایر الگوریتم‌های یادگیری عمیق مانند شبکهٔ حافظهٔ کوتاه‌مدت ماندگار (LSTM)، استفاده از روش‌های ترکیبی را پیشنهاد داده‌اند (Livieris and et al., 2021) که در پژوهش‌های آتی نیز می‌توان از ترکیب روش ارائه‌شده در این پژوهش با سایر الگوریتم‌های یادگیری عمیق استفاده کرد یا با ترکیب روش پیشنهادی با الگوریتم‌های فراابتکاری بهینه سازی مانند الگوریتم ازدحام ذرات (PSO)[4]و الگوریتم رقابت استعماری (ICA)[5]توانایی آنها را در قدرت پیش‌بینی قیمت مقایسه کرد.

از محدودیت‌های این نوشتار می‌توان به این نکته اشاره کرد که با توجه به وجود طیف وسیع شاخص‌های تکنیکال، متغیرهای تکنیکال استفاده شده در پژوهش حاضر نمایندهٔ کل شاخص‌های تکنیکال نیست. علاوه بر این، با در نظر گرفتن عمدهٔ پژوهش‌های گذشته در این حوزه، برای اجرای الگوریتم تنها از توابع فعال سازی ReLU و Tanh استفاده شده است؛ علاوه بر این، سیاست‌های مدیریتی شرکت‌ها و استفاده از معاملات بلوکی برای کنترل دستوری قیمت که بر حجم معاملات و قیمت کوتاه‌مدت اثرگذار است، ممکن است بر نتایج پژوهش اثرگذار باشد.

---

1. Principal Component Analysis
2. Singular Value Decomposition
3. Generalized Discriminant Analysis
4. Particle Swarm Optimization
5. Imperialist Competitive Algorithm

منابع فارسی

پورذاکر عربانی، سودابه. و ابراهیم پور کومله، حسین. (۱۳۹۸). بهینه سازی پیش‌بینی تقاضای وجه نقد دستگاه‌های خودپرداز شبکه بانکی کشور با استفاده از شبکهٔ عصبی بازگشتی عمیق LSTM. *تحقیق در عملیات در کاربردهای آن*، ۱۶(۳)، ۸۸-۶۹.

دامی، سینا. و خدابخش جویناني، فاطمه. (۱۳۹۶). پیش‌بینی بازار سهام با استفاده از یادگیری عمیق بازنمایی‌های عددی و متنی. *کنفرانس ملی دانش و فناوری نوین در علوم مهندسی در عصر تکنولوژی. تهران.*

راعی، رضا. و چاوشی، کاظم. (۱۳۸۲). پیش‌بینی بازده سهام در بورس اوراق بهادار تهران: مدل شبکهٔ عصبی مصنوعی و مدل چندعاملی. *تحقیقات مالی*، ۵ (۱۵)، ۱۲۰-۹۷.

شوشتریان، بیژن.، زاهدی، اسماعیل.، یزدی نژاد، محسن. و هرمزان، ساره. (۱۳۹۷). پیش‌بینی قیمت فلزات گران‌بها با استفاده از الگوریتم‌های یادگیری عمیق. *کنفرانس بین‌المللی نوآوری و تحقیق در علوم مهندسی. گرجستان – تفلیس.*

کیانی ماوی، رضا. و صیادی نیک، کامران. (۱۳۹۳) کاربرد الگوریتم‌های مختلف یادگیری در پیش‌بینی قیمت سهام با استفاده از شبکهٔ عصبی. *مدیریت توسعه و تحول*، ۱۳۹۳، ۷۵-۸۱

References

Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, 519, 127-139. https://doi.org/10.1016/j.physa.2018.11.061

Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of china stock market. *In 2015 IEEE International Conference on Big Data (Big Data)*, 2823-2824. https://doi.org/10.1109/BigData.2015.7364089

Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 83,187-205. https://doi.org/10.1016/j.eswa.2017.04.030

Dami, S., & Khodabaksh, J. F. (2017). Predicting the price of stock market using deep learning. *National Conference on Modern Knowledge and Technology in Engineering in the Technology Era.* Tehran. (In Persian)

Das, S. R., Mokashi, K., & Culkin, R. (2018). Are markets truly efficient? Experiments using deep learning algorithms for market movement prediction. *Algorithms*, 11(9), 138-157. https://doi.org/10.3390/a11090138

Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing*, 7(3–4), 197-387. http://dx.doi.org/10.1561/2000000039

Dezsi, E., & Nistor, I. A. (2016). Can deep machine learning outsmart the market? A comparison between econometric modelling and long- short term memory. *Romanian Economic Business Review*, 11(4.1), 54-73.

Gao, P., Zhang, R., & Yang, X. (2020). The application of stock index price prediction with neural network. *Mathematical and Computational Applications*, 25(3), 53-69. https://doi.org/10.3390/mca25030053

Gordon, J. A., & Francis, J. C. (1986). *Portfolio Analysis, (3th. Ed.).* California: Prentice-Hall.

Gudelek, M. U., Boluk, S. A., & Ozbayoglu, A. M. (2017). A deep learning based stock trading model with 2-D CNN trend detection. *In 2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1-8.

Gunduz, H., Yaslan, Y., & Cataltepe, Z. (2017). Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations. *Knowledge-Based Systems*, 137, 138-148. https://doi.org/10.1016/j.knosys.2017.09.023

Hiransha, M., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2018). Nse stock market prediction using deep-learning models. *Procedia Computer Science*, 132, 1351-1362. https://doi.org/10.1016/j.procs.2018.05.050

Hoseinzade, E., & Haratizadeh, S. (2019). CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, 129, 273-285. https://doi.org/10.1016/j.eswa.2019.03.029

Ji, L., Zou, Y., He, K., & Zhu, B. (2019). Carbon futures price forecasting based with ARIMA-CNN-LSTM model. *Procedia Computer Science*, 162, 33-38. https://doi.org/10.1016/j.procs.2019.11.254

Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert systems with Application*, 38 (5), 5311-5319. https://doi.org/10.1016/j.eswa.2010.10.027

Khare, K., Darekar, O., Gupta, P., & Attar, V. Z. (2017). Short term stock price prediction using deep learning. *In 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 482-486. https://www.doi.org/10.1109/RTEICT.2017.8256643

Kiyani Mavi, R., & Sayadi Nik, K. (2015). Using different learning algorithms in the stock price prediction by using neural networks. *Journal of Development and Evolution Management*, 1393, 75-81. (In Persian)

Lee, S. I., & Yoo, S. J. (2020). Threshold-based portfolio: The role of the threshold and its applications. *The Journal of Supercomputing*, 76(10), 8040-8057. https://doi.org/10.1007/s11227-018-2577-1

Li, Y., & Dai, W. (2020). Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model. *The Journal of Engineering*, 2020(13), 344-347. https://doi.org/10.1049/joe.2019.1203

Liu, S., Zhang, C., & Ma, J. (2017). CNN-LSTM neural network model for quantitative strategy analysis in stock markets. *International Conference on Neural Information Processing*, Springer, Cham. https://doi.org/10.1007/978-3-319-70096-0_21

Livieris, I. E., Kiriakidou, N., Stavroyiannis, S., & Pintelas, P. (2021). An advanced CNN-LSTM model for cryptocurrency forecasting. *Electronics*, 10(3), 287. https://doi.org/10.3390/electronics10030287

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Computing & Applications*, 32(23), 17351–17360. https://doi.org/10.1007/s00521-020-04867-x

Mitchell, T. (1997). *Machine Learning.* New York: McGraw Hill.

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications*, 42(4), 2162-2172. https://doi.org/10.1016/j.eswa.2014.10.031

Persio, L. D., & Honchar, O. (2016). Artificial neural networks architectures for stock price prediction: Comparisons and applications. *International Journal of Circuits, Systems and Signal Processing*, 10, 403-413. https://doi.org/10.1007/s40745-020-00305-w

Poorzaker Arabani, S., & Ebrahimpour Komleh, H. (2019). The optimization of forecasting ATMs cash demand of Iran banking network using LSTM deep recursive neural network. *Journal of Operational Research and Its Applications*, 16(3), 69-88. (In Persian)

Raei, R., & Chavoshi, K. (2003). Predicting the stock price: Artificial neural network and multi-factor model. *Financial Research Journal*, 15(5), 97-120. (In Persian)

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing Journal*, 90, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Sezer, O. B., & Ozbayoglu, A. M. (2018). Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. *Applied Soft Computing Journal*, 70, 525-538. https://doi.org/10.1016/j.asoc.2018.04.024

Shoushtarian, B., Zahedi, E., Yazdinejad, M., & Hormozan, S. (2018). Predicting the price of precious metal by deep learning algorithms. *International Conference on Innovation & Research on Engineering Sciences. Tbilisi.* (In Persian)

Zhou, F., Zhou, H. M., Yang, Z., & Yang, L. (2019). EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction. *Expert Systems with Applications*, 115, 136-151. https://doi.org/10.1016/j.eswa.2018.07.065

Zhou, X., Pan, Z., Hu, G., Tang, S., & Zhao, C. (2018). Stock market prediction on high frequency data using generative adversarial nets. *Mathematical Problems in Engineering*, 2018, 1–11.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی