ISSN Online: 2821-1936

Transactions on Data Analysis in Social Science

Journal Homepage: https://transoscience.ir

A Review of Deep Learning Methods for Financial Market Prediction

A. Karimi Dastgerdi[1], F. Zamani Boroujeni[2,*]

[1] Department of Computer Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran
[2] Department of Computer Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

ARTICLE INFO

Article History:
Received 4 June 2020
Received in revised form 18 August 2020
Accepted 20 September 2020
Available online 23 September 2020

Keywords:
Deep Learning, Financial Markets, Time Series Prediction, Stock Price, Stock Exchange

ABSTRACT

Financial market prediction plays a crucial role for investors and financial institutions aiming to optimize returns and minimize risks. Over the past decades, considerable research has focused on developing effective and accurate methods for forecasting financial market trends. Traditional statistical models often face limitations in capturing the nonlinear and complex dynamics of financial time series. In contrast, deep learning techniques provide advanced analytical and predictive frameworks capable of uncovering latent structures and intricate patterns within large-scale financial datasets. This study systematically reviews recent deep learning approaches applied to the prediction of financial time series, including recurrent neural networks (RNNs), long short-term memory networks (LSTMs), convolutional neural networks (CNNs), and hybrid models. We evaluate these methods based on their architecture, data representation, and predictive performance, highlighting their respective strengths and weaknesses. Our analysis demonstrates that deep learning algorithms exhibit superior capabilities in modeling nonlinear dependencies and temporal correlations in financial data, enabling accurate forecasting of stock prices, indices, and other market indicators. Furthermore, while individual network architectures perform effectively, combining recurrent and convolutional layers often enhances prediction accuracy and robustness. The findings underscore the potential of deep learning as a powerful tool for financial decision-making, offering valuable insights for both researchers and practitioners in the field of computational finance.

1. INTRODUCTION

In recent years, deep learning has attracted significant attention and has been widely discussed in the fields of artificial intelligence and machine learning. Deep learning refers to a subset of machine learning algorithms and methods that aim to uncover complex patterns in data and model high-level abstractions. This capability has led to remarkable successes across machine learning applications, even surpassing human performance in certain domains such as object recognition.

* Corresponding Author: f.zamani@khuisf.ac.ir
Department of Computer Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

http://dx.doi.org/10.47176/TDASS.2020.164

Some deep learning techniques focus on the analysis and prediction of time series data, assisting in forecasting by discovering hidden and previously unknown patterns. Given the nonlinear and chaotic nature of financial markets, traditional analytical approaches often lack sufficient accuracy. Therefore, this study explores and compares modern deep learning methods applied to financial market prediction.

2. METHODS AND RECENT DEVELOPMENTS

Since the early 1970s, and particularly from the 1980s onward, extensive efforts have been made to forecast stock prices using new mathematical methods, time series analysis, and more advanced tools such as artificial intelligence. Numerous experiments have been conducted on stock price and index data in countries including the United States, the United Kingdom, Canada, Germany, and Japan to examine whether specific patterns exist within price data and, consequently, to challenge the random walk hypothesis [1].

Proponents of chaos theory argue that beneath the seemingly random patterns of various phenomena ranging from meteorological systems to organizations and financial markets there exists an inherent order [2]. According to this theory, if such systems are observed over a sufficient period, one can discern that they consistently preserve an intrinsic structure. Even the most unpredictable systems move within defined boundaries and never entirely depart from them.

In recent years, a variety of deep learning methods have been employed to predict financial markets. Generally, these approaches can be classified into four main categories based on the structure of the deep neural network used:

1. Deep Neural Networks (DNNs)
2. Recurrent Neural Networks (RNNs)
3. Long Short-Term Memory Networks (LSTMs)
4. Convolutional Neural Networks (CNNs)

2.1. Deep Neural Networks (DNNs)

Deep Neural Networks (DNNs) are inspired by the structure and function of the human brain [3]. To train a DNN, researchers stack multiple layers of neural networks sequentially. Figure 1 illustrates a typical neuron, the fundamental building block of a deep neural network.

Although research on deep architectures has been ongoing for decades, no successful method for training deep networks was reported until 2006, primarily due to difficulties in training such architectures effectively. In that year, Hinton et al. introduced Deep Belief Networks (DBNs) and proposed a greedy layer-wise training algorithm, which trained each layer individually using unsupervised learning techniques [4].

These networks are data-driven and perform automatic feature extraction, leading to high accuracy and superior performance across diverse domains. Increasing the depth of the network enhances its representational capacity; however, challenges such as insufficient data, hardware limitations, overfitting, and the vanishing gradient problem initially limited their application. In recent years, the resolution of these issues has resulted in the widespread adoption and success of deep neural networks [5].

For the first time, White [6] utilized neural networks to predict financial markets. By using a case study on the daily stock prices of IBM, he demonstrated that neural networks are capable of identifying nonlinear rules in time series data as well as hidden relationships and stock price fluctuations. Kim and Han [7] employed a neural network optimized with a genetic algorithm to forecast stock indices and achieved promising results. Similarly, Garliauskas [8] predicted stock market time series using a kernel-based neural computation algorithm combined with an error backpropagation prediction method. His findings revealed that neural networks outperform classical statistical models and other traditional methods in forecasting financial time series.

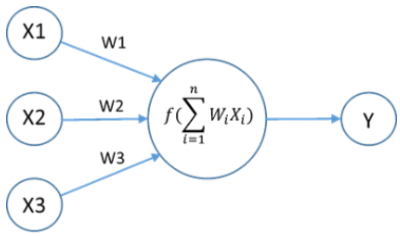

Fig. 1. A neuron as the computational unit of a deep neural network

In a 2017 study, Chong et al. [9] investigated the effect of three unsupervised feature extraction techniques Principal Component Analysis (PCA), Autoencoders, and Restricted Boltzmann Machines (RBMs) on the ability of deep neural networks to forecast future stock market behavior. They also sought to extract additional information from the residuals of autoregressive (AR) models and obtained encouraging results. Their findings showed that deep neural networks can extract additional information from autoregressive model residuals, thereby enhancing prediction performance. Conversely, autoregressive models could not capture supplementary information from the residuals of neural network models.

Furthermore, the study found that when the predictive network was applied to covariance-based market structures, it significantly improved covariance estimation. When comparing deep learning models with traditional linear autoregressive models, the results demonstrated that deep learning approaches yielded superior predictive performance. One limitation of their study, however, was the exclusion of trading volume and derivative prices related to stock prices from the input data.

2.2. Long Short-Term Memory Networks (LSTM)

A Long Short-Term Memory (LSTM) network is a type of recurrent neural network (RNN) that contains LSTM blocks instead of conventional recurrent units. An LSTM block can be described as an intelligent network unit capable of retaining information over arbitrary time intervals. Each block includes gates that determine when to store, forget, or output information [10].

The LSTM neural network architecture was introduced by Hochreiter and Schmidhuber in 1997 [11], and its structure is illustrated in Figure 3. Unlike traditional neural networks, an LSTM is particularly well-suited for learning from experiences involving classification, processing, and prediction of time series data especially when there are long and uncertain time lags between significant events. This property explains why LSTM networks often outperform standard RNNs, Hidden Markov Models (HMMs), and other sequential learning methods in various applications [12].

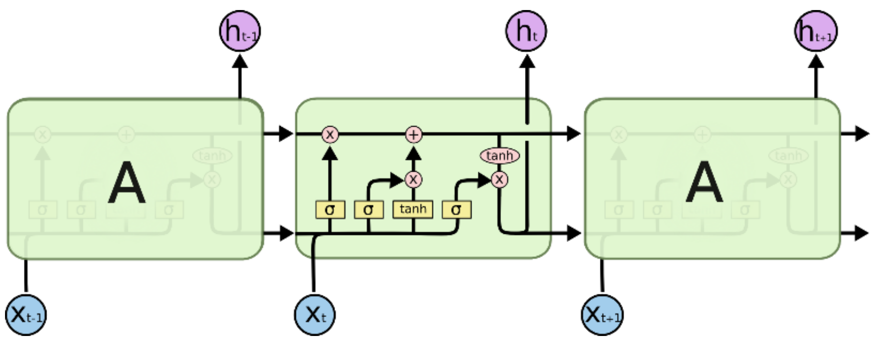

Fig. 2. Architecture of Long Short-Term Memory (LSTM) Networks

In 2017, Bao et al. [13] proposed a deep learning method for stock price prediction that integrated Long Short-Term Memory (LSTM) networks, wavelet transforms, and stacked autoencoders (SAEs). The proposed deep learning framework consisted of three stages. First, the stock price time series was decomposed using a wavelet transform to reduce noise. Second, a stacked autoencoder was employed to generate high-level features for stock price forecasting. Finally, the denoised high-level features were fed into an LSTM neural network to predict the next day's stock price.

One of the main advantages of their approach was the use of stacked autoencoders for high-level feature extraction, applied for the first time in stock price prediction. Another advantage was the inclusion of modular components for noise elimination. The authors compared their model with standard LSTM, RNN, and Wavelet-LSTM (WLSTM) models and found that their framework achieved higher predictive accuracy and profitability, regardless of the chosen stock index or input dataset.

Despite these advantages, their model suffered from non-optimized hyperparameters within the deep learning framework. The authors suggested that performance could be further improved through more sophisticated hyperparameter tuning strategies.

In 2018, Haijoong Chung et al. [14] introduced a hybrid approach that combined LSTM networks with a genetic algorithm (GA). Typically, the window size and architectural parameters of LSTM networks are determined through trial-and-error or heuristic search methods. However, the authors proposed a GA-based technique capable of automatically optimizing both the time window size and its position for LSTM networks.

Their experimental results demonstrated improved model performance compared to baseline LSTM models. To evaluate accuracy, they employed standard metrics including Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The authors emphasized that defining an appropriate time window is crucial for identifying temporal dependencies in financial datasets. Moreover, they claimed that their model overcame the limitations of traditional statistical methods, which often rely on restrictive statistical assumptions.

Another key advantage of their hybrid model was its ability to identify noisy patterns in financial time series data a property that could be extended to other application areas. However, the study had some limitations, such as excluding influential factors like transaction commissions from stock price modeling. Additionally, interpreting the outputs of LSTM networks remained challenging, a problem that could be mitigated by integrating other machine learning techniques.

In another 2018 study, Chongyi Song et al. [15] employed LSTM networks to enhance the performance of traditional technical analysis algorithms used in trading systems. They observed that most financial data exhibit strong time-series correlations. Their results showed that combining deep learning methods, such as LSTM, with traditional technical analysis produced superior forecasting accuracy compared to using either approach independently.

Similarly, in 2018, Fischer and Krauss [16] conducted a study using S&P 500 stock data to predict future stock price movements based on LSTM networks. This represented the first application of LSTM networks for forecasting stock market profitability. Comparing their LSTM model with traditional techniques, they found that the LSTM achieved higher predictive accuracy, and trading strategies derived from its predictions demonstrated the best performance, with a Sharpe ratio of 8.5 surpassing all other evaluated models. However, their model's performance declined for stock data recorded after 2010.

The authors also sought to gain insight into the "black box" of deep learning by examining common patterns identified in traded stocks. They discovered that most of the network's predictions were based on mean reversion behavior, offering an interpretive window into the model's internal decision-making processes.

2.3. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) represent a significant advancement achieved through deep layered architectures for information processing, particularly in the field of image analysis. Although the concept of CNNs was first developed in the 1980s and 1990s, the technological limitations of that era prevented their effective application in real-world image processing tasks, leading to a temporary decline in their use. Since their resurgence in 2012, CNNs have revolutionized the field of computer vision and have continued to grow at an accelerating pace. Figure 4 illustrates the architecture of a typical convolutional neural network.

A CNN is composed of three main layers: the convolutional layer, the pooling layer, and the fully connected layer. Each of these layers performs a distinct function. The convolutional layer generates new images known as feature maps, which highlight specific characteristics of the original image. Unlike traditional neural network layers, the convolutional layer does not rely on connection weights and weighted summations. Instead, it employs filters (or kernels) that transform input images to extract local features.

The pooling layer reduces the spatial dimensions of the image by combining neighboring pixels within a specific region into a representative value typically the mean or maximum of those pixels. Pooling is a common technique that has also been widely used in earlier image processing methods [17].

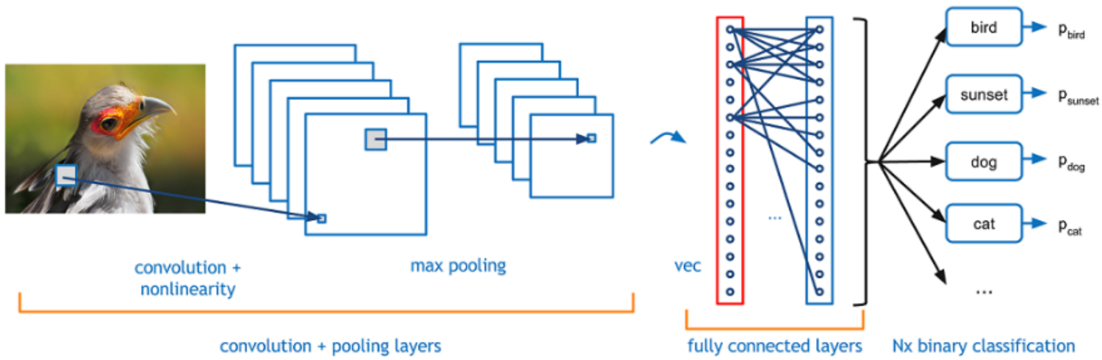

Fig. 3. An Overview of Convolutional Neural Network Architecture

Selvin et al. [18] (2017) proposed a moving-window-based approach capable of identifying underlying dynamics and interrelationships within data through deep learning architectures, which they applied to short-term stock price forecasting. They implemented and compared three deep learning architectures Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Convolutional Neural Networks (CNNs) for predicting the Indian stock market. Their findings revealed that CNNs are particularly effective in detecting shifts in stock market trends and outperform the other two models in predictive accuracy.

Avram Tsantekidis et al. [19] (2017) introduced a deep learning method based on CNNs that predicts stock price movements using large-scale, high-frequency time-series data derived from financial market order books. They compared their CNN-based model with Multilayer Perceptrons (MLPs) and Support Vector Machines (SVMs). The results demonstrated that their approach achieved superior performance in short-term stock movement prediction compared to traditional models. However, the method’s limitations include the lack of noise removal from input data, which may significantly distort prediction outcomes, and the absence of advanced normalization techniques, potentially reducing prediction accuracy.

Hiransha et al. [20] (2018) conducted a comparative study using the ARIMA algorithm and four deep learning models MLP, LSTM, RNN, and CNN for predicting the Indian stock market. Their models were trained using New York Stock Exchange (NYSE) data and subsequently tested on Indian stock data to assess generalization capability. The results showed that while ARIMA effectively learns hidden patterns within a specific time series, it performs poorly on different datasets. In contrast, deep neural networks demonstrated the ability to generalize and accurately predict the Indian stock market despite being trained on NYSE data. Among all tested models, CNNs achieved the highest predictive accuracy, effectively identifying sudden fluctuations in time-series data.

Hyun-Seok Sim et al. [21] (2019) explored the applicability of CNNs traditionally used for image classification to stock market prediction. In their study, they converted technical indicators of stock prices into graphical time-series images and used these as input variables. They compared the predictive accuracy of their proposed CNN model with SVMs and ANNs, concluding that CNNs are highly suitable for financial time-series forecasting. A major strength of their approach lies in the optimization of CNN parameters. However, their findings indicated that technical indicators alone do not significantly enhance model performance. The study’s limitations include the exclusion of key macroeconomic factors such as gold and currency prices and the effects of inflation, which can strongly influence stock movements.

Wenlong et al. [22] (2019) proposed an end-to-end model called the Multi-Filter Neural Network (MFNN) for feature extraction from financial time-series data and stock price movement prediction. They integrated CNNs and RNNs to capture spatial and temporal dependencies across multiple feature positions, thereby developing a multi-filter architecture. Experimental results demonstrated that their model achieved superior accuracy, profitability, and stability compared to traditional machine learning, statistical models, and single-architecture deep learning models (such as LSTM, RNN, and CNN). However, their method still suffered from risk instability, which they suggested could be mitigated by better understanding the effect of filter integration on trading signal quality.

Xin-Yu Zhou et al. [23] (2018) combined LSTM and CNN architectures within an adversarial training framework to predict high-frequency stock markets. They used a set of thirteen common technical indicators available in trading software as input features. One of the key advantages of their method was the use of a rolling training window, enabling them to analyze how the model update cycle affects predictive performance. They compared their approach with ARIMA-GARCH, ANN, SVM, and LSTM models, finding that their proposed method significantly outperformed all others. Their results showed that shorter update cycles improve forecast accuracy, and adversarial training effectively reduced both prediction costs and forecast errors. They named their model GAN-FD and suggested that its ease of use could assist non-expert investors in making informed financial decisions.

3. SUMMARY AND CONCLUSION

The deep learning methods reviewed in this study for financial market prediction are summarized in **Table 1**, which also presents the datasets used for model training, the input variables considered, and the accuracy evaluation metrics employed.

Table 1. Comparison of Proposed Methods for Financial Market Prediction

Authors	Dataset	Input Variables	Proposed Method	Accuracy Metric
Weibao et al. (2017)	CSI300, Nifty50, Hong Kong Market, Nikkei225, S&P500	Stock Prices	WT + SAEs + LSTM	MAPE, R, Theil's U
Chung et al. (2017)	Korea KOSPI	Profitability	Data Representation + DNN	NMSE, RMSE, MAE, MI
Selvin et al. (2017)	NSE	Stock Prices and Trading Volume	CNN	Percentage Error
Tsantekidis et al. (2017)	S&P500	Limit Order Book Data	CNN	Cohen's Kappa, Mean Recall, Precision, F1
Haijun Chung et al. (2018)	Korea KOSPI	Stock Prices, Trading Volume	GA + LSTM	MSE, MAE, MAPE
Wenlong et al. (2018)	CSI 300	Stock Prices, Trading Volume	LSTM + CNN	Total Return
Xin-Yu Zhou et al. (2018)	CSI 300	Stock Prices	GANs + LSTM + CNN	RMSRE, DPA
Chengji Song et al. (2018)	S&P500	Technical Indicators	LSTM	Cross-Entropy Cost
Fischer & Krauss (2018)	S&P500	Return Rate	LSTM	R^2, RMSE
Hiransha et al. (2018)	NSE, NYSE	Stock Prices	CNN	MAPE
Hyun-Seok Sim et al. (2019)	S&P500	Closing Prices and Technical Indicators	CNN	MSE, CEE

As demonstrated, deep learning algorithms have made a remarkable impact on improving financial market forecasting methods. While recurrent neural networks (RNNs) are inherently designed for time-series analysis, convolutional neural networks (CNNs) can also be effectively employed to detect hidden patterns within time-series data, achieving high prediction accuracy.

Based on the findings of previous research, it can be concluded that although both Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNNs) independently yield strong performance, their hybrid integration further enhances both accuracy and efficiency in financial market prediction tasks.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFERENCES

[1] Pinches, G. E. (1970). The random walk hypothesis and technical analysis. Financial Analysts Journal, 26(2), 104–110. https://doi.org/10.2469/faj.v26.n2.104

[2] Weiss, G. (1992). Chaos hits Wall Street the theory that is! Business Week, 2, 138–140.

[3] Hrasko, R., Pacheco, A. G., & Krohling, R. A. (2015). Time series prediction using restricted Boltzmann machines and backpropagation. Procedia Computer Science, 55, 990–999. https://doi.org/10.1016/j.procs.2015.07.104

[4] Ying, C., Yang, M.-L., & Jun, L. (2015). Multiclass classification based on a deep convolutional neural network. Frontiers of Information Technology & Electronic Engineering, 16(11), 930–939. https://doi.org/10.1631/FITEE.1500125

[5] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. Neurocomputing, 234, 11–26. https://doi.org/10.1016/j.neucom.2016.12.038

[6] White, H. (1988). Economic prediction using neural networks: The case of IBM daily stock returns. In Proceedings of the IEEE International Conference on Neural Networks (pp. 451–458). IEEE. https://doi.org/10.1109/ICNN.1988.23959

[7] Kim, K.-J., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. Expert Systems with Applications, 19(2), 125–132. https://doi.org/10.1016/S0957-4174(00)00027-0

[8] Garliauskas, A. (1999). Neural network chaos and computational algorithms of forecast in finance. In Proceedings of the 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028) (Vol. 2, pp. 638–643). IEEE. https://doi.org/10.1109/ICSMC.1999.825335

[9] 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

[10] Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2), 157–166. https://doi.org/10.1109/72.279181

[11] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[12] Graves, A., Mohamed, A.-R., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 6645–6649). IEEE. https://doi.org/10.1109/ICASSP.2013.6638947

[13] Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long short-term memory. PLOS ONE, 12(7), e0180944. https://doi.org/10.1371/journal.pone.0180944

[14] Chung, H., & Shin, K.-S. (2018). Genetic algorithm-optimized long short-term memory network for stock market prediction. Sustainability, 10(10), 3765. https://doi.org/10.3390/su10103765

[15] Sang, C., & Di Pierro, M. (2019). Improving trading technical analysis with TensorFlow long short-term memory (LSTM) neural network. The Journal of Finance and Data Science, 5(1), 1–11. https://doi.org/10.1016/j.jfds.2018.10.003

[16] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

[17] Kim, P. (2017). Matlab deep learning: With machine learning, neural networks and artificial intelligence. Springer.

[18] Selvin, S., Vinayakumar, R., Gopalakrishnan, E., Menon, V. K., & Soman, K. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1643–1647). IEEE. https://doi.org/10.1109/ICACCI.2017.8126078

[19] Tsantekidis, A., Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., & Iosifidis, A. (2017). Forecasting stock prices from the limit order book using convolutional neural networks. In 2017 IEEE 19th Conference on Business Informatics (CBI) (Vol. 1, pp. 7–12). IEEE. https://doi.org/10.1109/CBI.2017.23

[20] Hiransha, M., Gopalakrishnan, E., Menon, V. K., & Soman, K. (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

[21] Sim, H. S., Kim, H. I., & Ahn, J. J. (2019). Is deep learning for image recognition applicable to stock market prediction? Complexity, 2019, Article 4324878. https://doi.org/10.1155/2019/4324878

[22] Long, W., Lu, Z., & Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. Knowledge-Based Systems, 164, 163–173. https://doi.org/10.1016/j.knosys.2018.10.034

[23] 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, Article 4907423. https://doi.org/10.1155/2018/4907423