Stock Market Prediction Using Deep Reinforcement Learning

Subjects: Computer Science, Artificial Intelligence

Contributor: Alamir Labib Awad , Saleh Mesbah Elkaffas , Mohammed Waleed Fakhr

Stock market investment, a cornerstone of global business, has experienced unprecedented growth, becoming a lucrative, yet complex field. Predictive models, powered by cutting-edge technologies like artificial intelligence (AI), sentiment analysis, and machine learning algorithms, have emerged to guide investors in their decision-making processes.

stock trading markets deep reinforcement learning DRL neural networks stock prediction

1. Introduction

Stock market investment, a cornerstone of global business, has experienced unprecedented growth, becoming a lucrative, yet complex field [1][2]. Predictive models, powered by cutting-edge technologies like artificial intelligence (AI), sentiment analysis, and machine learning algorithms, have emerged to guide investors in their decision-making processes [3][4][5]. Key among these techniques are convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM), all rooted in neural network methodologies. These intelligent software systems assist traders and investors in augmenting their trading strategies [6]. However, existing predictive models struggle to adapt swiftly to unforeseen market events, influenced by intricate external factors such as economic trends, market dynamics, firm growth, consumer prices, and industry-specific shifts. These factors impact stock prices, leading to unpredictable outcomes [7][8]. Hence, a fundamental analysis integrating economic factors and the ability to analyze financial news and events is imperative. Historical datasets, fundamental to stock models, often contain noisy data, demanding meticulous handling for accurate predictions. The volatile nature of stock markets, characterized by rapid fluctuations, requires precise predictions [9][10]. Diverse sources of stock market data, including media, news headlines, articles, and tweets, play a crucial role. Natural language processing (NLP) algorithms, particularly sentiment analysis, enable the extraction of sentiments from social media, news feeds, or emails. Sentiments are categorized as positive, negative, or neutral through machine learning (ML) or deep learning (DL) algorithms.

2. Stock Market Prediction Using Deep Reinforcement Learning

Stock price prediction efforts have centered on supervised learning techniques, such as neural networks, random forests, and regression methods [11]. A detailed analysis by authors [12] underscored the dependency of supervised models on historical data, revealing constraints that often lead to inaccurate predictions. In a separate study [13], speech and deep learning (DL) techniques were applied to stock prediction using Google stock datasets from NASDAQ. The research demonstrated that employing 2D principal component analysis (PCA) with deep neural networks (DNN) outperformed the results obtained with two-directional PCA combined with radial bias function neural network (RBFNN), highlighting the efficacy of specific methodologies in enhancing accuracy. Another comprehensive survey [14] explored various DL methods, including CNN, LSTM, DNN, RNN, RL, and others, in conjunction with natural language processing (NLP) and WaveNet. Utilizing datasets sourced from foreign exchange stocks in Forex markets, the study employed metrics like mean absolute percentage error (MAPE), root mean square error (RMSE), mean square error (MSE), and the Sharpe ratio to evaluate performance. The findings highlighted the prominence of RL and DNN in stock prediction research, indicating the increasing popularity of these methods in financial modeling. While this study covered a wide array of prediction techniques, it notably emphasized the absence of results related to combining multiple DL methods for stock prediction. In a different studies [15][16], four DL models utilizing data from NYSE and NSE markets were examined: MLP, RNN, CNN, and LSTM. These models, when trained separately, identified trend patterns in stock markets, providing insights into shared dynamics between the two stock markets. Notably, the CNN-based model exhibited superior results in predicting stock prices for specific businesses. However, this study did not explore hybrid networks, leaving unexplored potential in creating combined models for stock prediction. Additionally, advances in machine learning have led to considerable progress in speech recognition, language processing, and image classification across various applications [17]. Researchers have applied digital signal processing methods to stock data, particularly focusing on time series data analysis [18]. Moreover, reinforcement learning (RL) has emerged as a method capable of overcoming the limitations of traditional supervised learning approaches. By combining financial asset price prediction with the allocation step, RL algorithms can make optimal decisions in the complex stock market environment [19]. While LSTM techniques have been extensively researched for stock prediction due to their ability to efficiently process large datasets, challenges arise from the need for substantial historical data and considerable computational resources [20]. A critical issue with LSTM models is their limited capacity to offer rational decisions to investors, such as whether to buy, sell, or retain stocks based on predictions [21]. However, a recent study [22] demonstrated the potential of combining LSTM with sentiment analysis, providing valuable support to stock investors in decision-making processes. Furthermore, researchers have explored support vector machine (SVM) techniques in time series prediction. Despite their accuracy, SVM models require extensive datasets and involve time-consuming training processes [23]. In the comprehensive review of existing literature, it became evident that both supervised and unsupervised machine learning models have limitations, despite their efficiency in predicting time series data. Researchers have identified specific challenges associated with raw data characteristics, leading to barriers to accurate stock market predictions [24][25].

3. Background

3.1. Deep Learning

Artificial neural networks (ANNs) replicate the complex operations of the human brain, enabling tasks such as classification and regression. ANNs comprise interconnected neurons organized in layers. Traditionally limited to a few layers due to computational constraints, modern ANNs, powered by GPUs and TPUs, support numerous hidden layers, enhancing their ability to detect nonlinear patterns as shown in **Figure 1**. Deep learning with ANNs finds applications in diverse fields, including computer vision, health care, and predictive analysis.

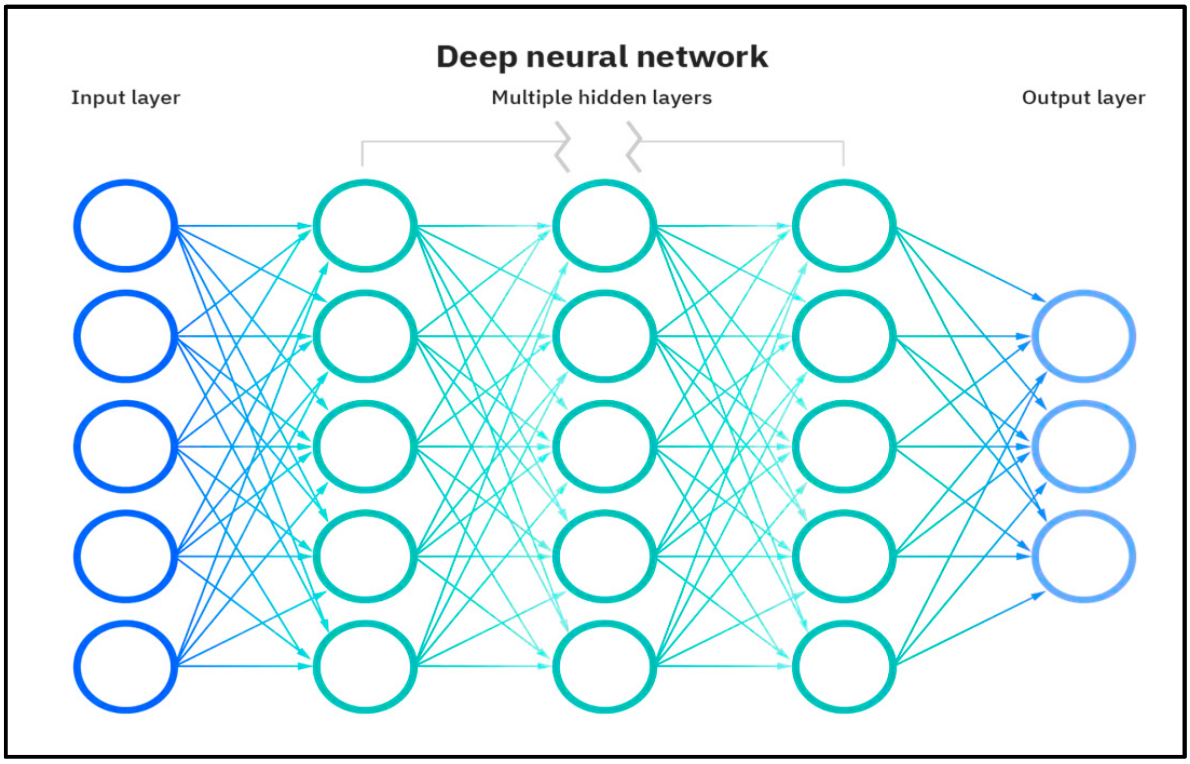


Figure 1. The architecture of an artificial neural network.

3.2. Recurrent Neural Network

Recurrent neural networks (RNNs) excel in processing sequential data. They possess a memory feature, retaining information from previous steps in a sequence as shown in **Figure 2**. RNNs incorporate inputs (“x”), outputs (“h”), and hidden neurons (“A”). A self-loop on hidden neurons signifies input from the previous time step (“t − 1”). However, RNNs face challenges like the vanishing gradient problem, mitigated by techniques like long short-term memory (LSTM) units. For instance, if the input sequence comprises six days of stock opening price data, the network unfurls into six layers, each corresponding to the opening stock price of a single day. However, a significant challenge confronting RNNs is the vanishing gradient problem, which has been effectively addressed through various techniques, including the incorporation of long short-term memory (LSTM) units into the network.

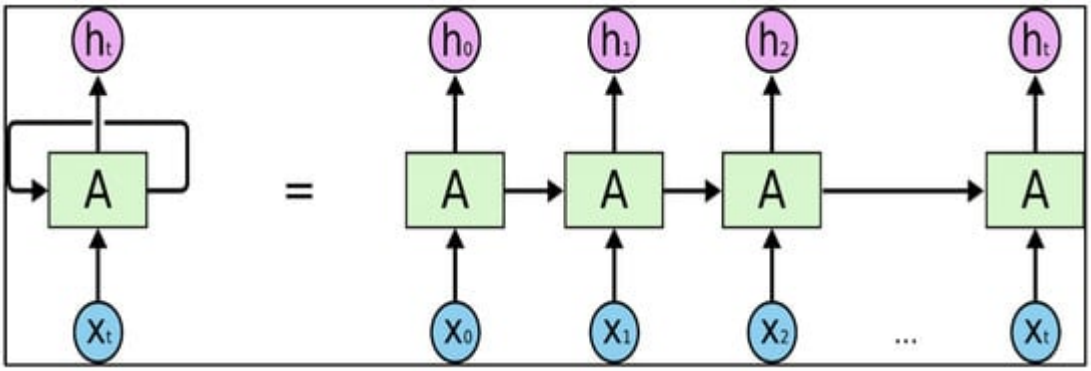


Figure 2. Unfolded recurrent neural network.

3.3. LSTM

LSTM enhances RNNs' memory, crucial for handling sequential financial data. LSTM units, integrated into RNNs, have three gates: input gate (i), forget gate (f), and output gate (o). These gates use sigmoid functions to write, delete, and read information, addressing long-term dependencies and preserving data patterns. In the LSTM architecture illustrated in **Figure 3**, three gates play pivotal roles:

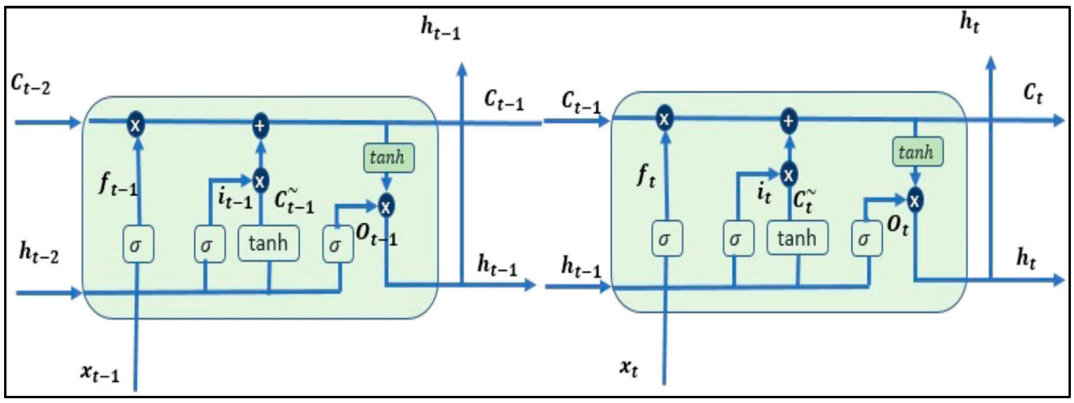


Figure 3. LSTM architecture.

- Input Gate (i): This gate facilitates the addition of new information to the cell state.
- Forget Gate (f): The forget gate selectively discards information that is no longer relevant or required by the model.
- Output Gate (o): Responsible for choosing the information to be presented as the output.

Each of these gates operates utilizing sigmoid functions, transforming values into a range from zero to one. This mechanism empowers LSTMs to adeptly write, delete, and read information from their memory, rendering them exceptionally skilled at handling long-term dependencies and preserving crucial patterns in data. Crucially, LSTMs address the challenge of the vanishing gradient, ensuring that gradient values remain steep enough during training.

This characteristic significantly reduces training times and markedly enhances accuracy, establishing LSTMs as a foundational technology in the domain of sequence prediction, especially for intricate datasets prevalent in financial markets.

3.4. Reinforcement Learning

Reinforcement learning involves an agent making decisions in different scenarios. It comprises the agent, environment, actions, rewards, and observations. Reinforcement learning faces challenges such as excessive reinforcements and high computational costs, especially for complex problems. The dynamics of reinforcement learning are encapsulated in **Figure 4**, illustrating the interaction between the agent and its environment. Notably, states in this framework are stochastic, meaning the agent remains unaware of the subsequent state, even when repeating the same action.

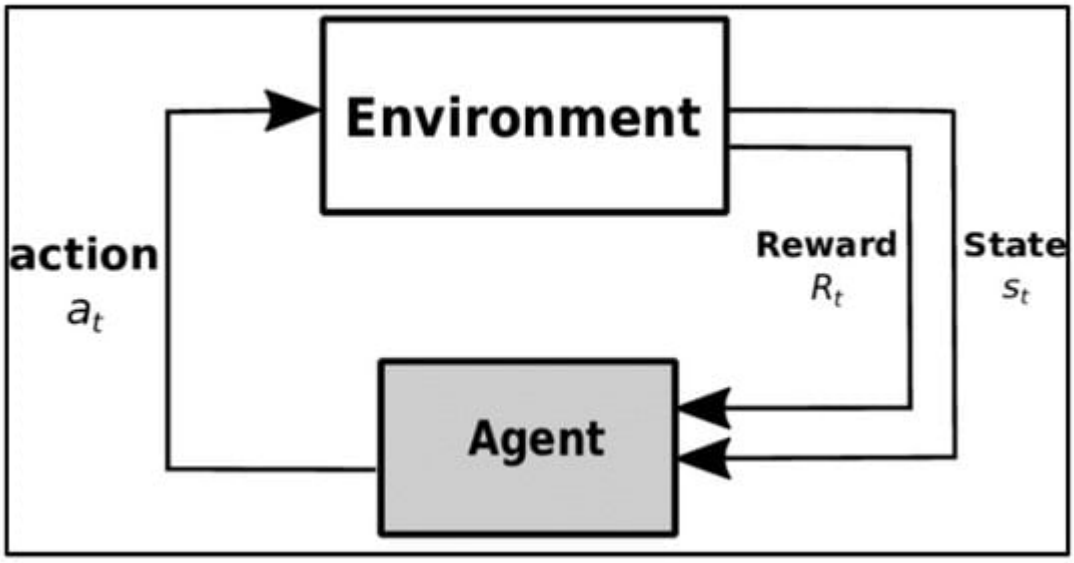


Figure 4. The reinforcement learning process.

Within the realm of reinforcement learning, several crucial quantities are determined:

- Reward: A scalar value from the environment that evaluates the preceding action. Rewards can be positive or negative, contingent upon the nature of the environment and the agent's action.
- Policy: This guides the agent in deciding the subsequent action based on the current state, helping the agent navigate its actions effectively.
- Value (V): Represents the long-term return, factoring in discount rates, rather than focusing solely on short-term rewards (R).
- Action Value: Like the reward value, but incorporates additional parameters from the current action. This metric guides the agent in optimizing its actions within the given environment.

Despite the advantages of reinforcement learning over supervised learning models, it does come with certain drawbacks. These challenges include issues related to excessive reinforcements, which can lead to erroneous outcomes. Additionally, reinforcement learning methods are primarily employed for solving intricate problems, requiring substantial volumes of data and significant computational resources. The maintenance costs associated with this approach are also notably high.

This study focuses on predicting gold prices based on next-day tweets sourced from news and media datasets. Gold prices exhibit rapid fluctuations daily, necessitating a robust prediction strategy. To achieve accurate predictions, this research employs a comprehensive approach integrating deep reinforcement learning (DRL), long short-term memory (LSTM), variational mode decomposition (VMD), and natural language processing (NLP). The prediction time spans from 2012 to 2019, utilizing tweets related to gold prices. DRL is enhanced by incorporating sentiment analysis of media news feeds and Twitter data, elevating prediction accuracy. The dataset used for this analysis was retrieved from the link https://www.kaggle.com/datasets/ankurzing/sentiment-analysis-in-commodity-market-gold accessed on 1 February 2023. This dataset, spanning from 2000 to 2021, encompasses diverse news sources and is meticulously classified as positive or negative by financial experts, ensuring the robustness and reliability of the data.

3.5. Deep Reinforcement Learning

Reinforcement learning (RL) operates as a trial-and-error methodology aimed at maximizing desired outcomes. Deep reinforcement learning (DRL) combines principles of deep learning and RL, where neural networks are trained to generate values crucial for reinforcement learning, as illustrated in **Figure 5**. DRL leverages prior learning from the environment and applies this knowledge to new datasets, enhancing its adaptability and learning capabilities. This approach revolves around a value function, defining the actions undertaken by the agent. In the realm of RL, the state is inherently stochastic, mirroring the inherent randomness and transitions found in variables within dynamic environments like stock markets. These variables shift between states based on underlying assumptions and probabilistic rules [26][27]. The Markov decision process (MDP) serves as a fundamental framework for modeling stochastic processes involving random variables. MDPs are instrumental in describing RL problems, particularly in managing tasks within rapidly changing environments [28]. Within the RL framework, the agent, functioning as a learner or decision-maker, interacts with the environment. In the context of MDP, the interactions between the agent and the environment define the learning process. At each step, denoted as $t \in \{1, 2, 3, \ldots, T\}$, the agent receives information about the current state of the environment, represented as $s_t \in S$. Based on this information, the agent selects and executes an action, denoted as $a_t \in A$. Subsequently, if the agent transitions to a new state, the environment provides a reward, $R_{(t + 1)} \in R$, to the agent as feedback, influencing the quality of future actions. This iterative process encapsulates the essence of MDPs in RL problem-solving, forming a crucial foundation for adaptive learning strategies.

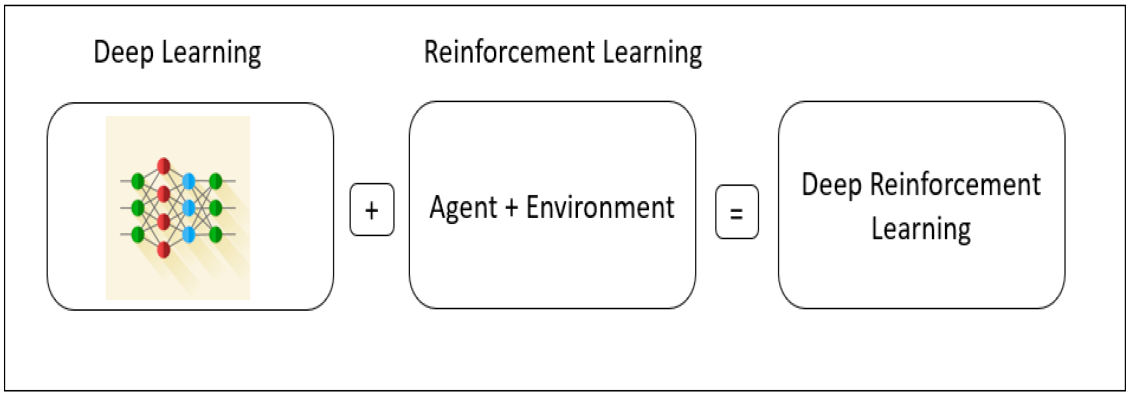


Figure 5. The DRL process.

3.6. Classification of the DRL Algorithms

Learning in DRL is based on actor or action learning, where policy learning is done to perform the best action at each state. The policy is obtained from data, and this learning continues with actions based on the learned policy. The agent will be trained in reinforcement learning based on critic-only, actor-only, and critic–actor approaches. RL algorithms are classified based on these three approaches [29].

In the critic-only approach, the algorithm will learn to estimate the value function by using a method known as generalized policy iteration (GPI). GPI involves the steps of policy evaluation, i.e., determining how good a given policy is and the next step of policy improvement. Here, the policy is improved by selecting greedy actions in relation to value functions obtained from the evaluation step. In this manner, the optimal policy is achieved [30].

3.7. Natural Language Processing

Natural language processing (NLP) analyzes natural languages such as English, French, etc., and makes computer systems interpret texts like humans. The human language is complicated to understand; hence, this is an ever-evolving field with endless applications. Every sentence should pass a preprocessing phase with six steps to build any NLP model. First is the tokenization phase, in which the sentence is split into a group of words. Second, the lowercasing phase converts every word to its lowercase form. Third, the stop words do not impact the sentence's meaning, so they are removed in this step. Fourth, every word is transformed into its root word in the steaming phase. Last, the lemmatization phase reduces the number of characters representing the word. After this preprocessing phase, there is the feature extraction in which the sentence is transformed from its textual representation into a mathematical representation called word embedding. Many word embedding approaches have been developed over the years. The classical approaches involve wrod2vec and Glove, while the modern ones include BERT.

3.8. Sentiment Analysis

Sentiment analysis aims to identify the opinion toward a product from a text. There are three modes toward a product: positive, negative, and neutral. Two main approaches are used in sentiment analysis: the supervised approach and the lexicon approach. In the supervised approach, the sentences are provided to the classification model along with their label, positive or negative. Then, the sentences are transformed into vectors, and the model makes a classification for these vectors.

On the other hand, the lexicon-based approach relies on the language dictionary itself. The model has a list of positive and negative words. The sentences are divided into words, each with a semantic score. Finally, the model calculates the total semantics of the sentence and decides whether it is a positive or negative sentence.

3.9. TFIDF

TF-IDF stands for term frequency–inverse document frequency. It is used for document search by getting a query as input and finding the relevant documents as output. It is a statistical analysis technique used to know the importance of a word inside a document. It calculates the frequency of a word inside a document, compares it with the frequency of the word inside all documents, and compares the two values. The assumption is that if the word is repeated many times in a document and rarely appears in other documents, this means that this word is vital for this document.

3.10. BERT

Bidirectional encoder representations from transformers (BERT) is based on deep learning transformers for natural language processing. BERT is trained bidirectionally, which means it analyzes the word and the surrounding words in both directions. Reading in both directions allows the model to understand the context deeply. BERT models are already pretrained, so they already know the word representation and the relationships between them. BERT is a generic model that can be fine-tuned for specific tasks like sentiment analysis tasks. BERT contains a stack of transformers, each consisting of an encoder and decoder network. It has two versions, the base version and the large one, which gives the best results compared to any other model.

References

1. Idrees, S.M.; Alam, M.A.; Agarwal, P. A Prediction Approach for Stock Market Volatility Based on Time Series Data. IEEE Accesss 2019, 7, 17287–17298.
2. Bouteska, A.; Regaieg, B. Loss aversion, the overconfidence of investors and their impact on market performance evidence from the US stock markets. J. Econ. Financ. Adm. Sci. 2020, 25, 451–478.
3. Feng, F.; He, X.; Wang, X.; Luo, C.; Liu, Y.; Chua, T.S. Temporal Relational Ranking for Stock Prediction|ACM Transactions on Information Systems. ACM Trans. Inf. Syst. (TOIS) 2019, 37, 1–

30.

4. Dirman, A. Financial distress: The impacts of profitability, liquidity, leverage, firm size, and free cash flow. Int. J. Bus. Econ. Law 2020, 22, 17–25.
5. Ghimire, A.; Thapa, S.; Jha, A.K.; Adhikari, S.; Kumar, A. Accelerating Business Growth with Big Data and Artificial Intelligence. In Proceedings of the 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 7–9 October 2020.
6. Kurani, A.; Doshi, P.; Vakharia, A.; Shah, M. A Comprehensive Comparative Study of Artificial Neural Networks (ANN) and Support Vector Machines (SVM) on Stock Forecasting. Ann. Data Sci. 2021, 10, 183–208.
7. Beg, M.O.; Awan, M.N.; Ali, S.S. Algorithmic Machine Learning for Prediction of Stock Prices. In FinTech as a Disruptive Technology for Financial Institutions; IGI Global: Hershey, PA, USA, 2019; pp. 142–169.
8. Shah, D.; Isah, H.; Zulkernine, F. Stock Market Analysis: A Review and Taxonomy of Prediction Techniques. Int. J. Financ. Stud. 2019, 7, 26.
9. Yadav, A.; Chakraborty, A. Investor Sentiment and Stock Market Returns Evidence from the Indian Market. Purushartha-J. Manag. Ethics Spiritual. 2022, 15, 79–93.
10. Chauhan, L.; Alberg, J.; Lipton, Z. Uncertainty-Aware Lookahead Factor Models for Quantitative Investing. In Proceedings of the 37th International Conference on Machine Learning (PMLR), Virtual, 13–18 July 2020; Volume 119, pp. 1489–1499.
11. Nti, I.K.; Adekoya, A.F.; Weyori, B.A. A novel multi-source information-fusion predictive framework based on deep neural networks for accuracy enhancement in stock market prediction. J. Big Data 2021, 8, 17.
12. Sakhare, N.N.; Imambi, S.S. Performance analysis of regression-based machine learning techniques for prediction of stock market movement. Int. J. Recent Technol. Eng. 2019, 7, 655–662.
13. Singh, R.; Srivastava, S. Stock prediction using deep learning. Multimed. Tools Appl. 2016, 76, 18569–18584.
14. Hu, Z.; Zhao, Y.; Khushi, M. A Survey of Forex and Stock Price Prediction Using Deep Learning. Appl. Syst. Innov. 2021, 4, 9.
15. Hiransha, M.; Gopalakrishnan, E.A.; Menon, V.K.; Soman, K.P. NSE Stock Market Prediction Using Deep-Learning Models. Procedia Comput. Sci. 2018, 132, 1351–1362.
16. Patel, R.; Choudhary, V.; Saxena, D.; Singh, A.K. Review of Stock Prediction using machine learning techniques. In Proceedings of the 5th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 3–5 June 2021; pp. 840–847.

17. Kamath, U.; Liu, J.; Whitaker, J. Deep Learning for NLP and Speech Recognition; Springer: Cham, Switzerland, 2019; pp. 575–613.
18. Manolakis, D.; Bosowski, N.; Ingle, V.K. Count Time-Series Analysis: A Signal Processing Perspective. IEEE Signal Process. Mag. 2019, 36, 64–81.
19. Kabbani, T.; Duman, E. Deep Reinforcement Learning Approach for Trading Automation in the Stock Market. IEEE Access 2022, 10, 93564–93574.
20. Moghar, A.; Hamiche, M. Stock Market Prediction Using LSTM Recurrent Neural Network. Procedia Comput. Sci. 2020, 170, 1168–1173.
21. Ren, Y.; Liao, F.; Gong, Y. Impact of News on the Trend of Stock Price Change: An Analysis based on the Deep Bidirectional LSTM Model. Procedia Comput. Sci. 2020, 174, 128–140.
22. Jin, Z.; Yang, Y.; Liu, Y. Stock closing price prediction based on sentiment analysis and LSTM. Neural Comput. Appl. 2019, 32, 9713–9729.
23. Parray, I.R.; Khurana, S.S.; Kumar, M.; Altalbe, A.A. Time series data analysis of stock price movement using machine learning techniques. Soft Comput. 2020, 24, 16509–16517.
24. Duan, G.; Lin, M.; Wang, H.; Xu, Z. Deep Neural Networks for Stock Price Prediction. In Proceedings of the 14th International Conference on Computer Research and Development (ICCRD), Shenzhen, China, 7–9 January 2022.
25. Huang, J.; Liu, J. Using social media mining technology to improve stock price forecast accuracy. J. Forecast. 2019, 39, 104–116.
26. Iqbal, S.; Sha, F. Actor-Attention-Critic for Multi-Agent Reinforcement Learning. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 2961–2970.
27. Singh, V.; Chen, S.-S.; Singhania, M.; Nanavati, B.; Kar, A.K.; Gupta, A. How are reinforcement learning and deep learning algorithms used for big data-based decision making in financial industries—A review and research agenda. Int. J. Inf. Manag. Data Insights 2022, 2, 100094.
28. Padakandla, S. A survey of reinforcement learning algorithms for dynamically varying environments. ACM Comput. Surv. (CSUR) 2021, 54, 1–25.
29. Xiao, Y.; Lyu, X.; Amato, C. Local Advantage Actor-Critic for Robust Multi-Agent Deep Reinforcement Learning. In Proceedings of the International Symposium on Multi-Robot and Multi-Agent Systems (MRS), Cambridge, UK, 4–5 November 2021.
30. Ren, Y.; Duan, J.; Li, S.E.; Guan, Y.; Sun, Q. Improving Generalization of Reinforcement Learning with Minimax Distributional Soft Actor-Critic. In Proceedings of the IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece, 20–23 September 2020.

Retrieved from https://encyclopedia.pub/entry/history/show/117216