

Stock Market Prediction Using Deep Reinforcement Learning

Chennakesavan J, Prasanna S
Department of Computer Application – PG
(VISTAS)

Abstract: Stock value prediction and trading, a captivating and complex research domain, continues to draw heightened attention. Ensuring profitable returns in stock market investments demands precise and timely decision-making. The evolution of technology has introduced advanced predictive algorithms, reshaping investment strategies. Essential to this transformation is the profound reliance on historical data analysis, driving the automation of decisions, particularly in individual stock contexts. Recent strides in deep reinforcement learning algorithms have emerged as a focal point for researchers, offering promising avenues in stock market predictions. In contrast to prevailing models rooted in artificial neural network (ANN) and long short-term memory (LSTM) algorithms, this study introduces a pioneering approach. By integrating ANN, LSTM, and natural language processing (NLP) techniques with the deep Q network (DQN), this research crafts a novel architecture tailored specifically for stock market prediction. At its core, this innovative framework harnesses the wealth of historical stock data, with a keen focus on gold stocks. Augmented by the insightful analysis of social media data, including platforms such as S&P, Yahoo, NASDAQ, and various gold market related channels, this study gains depth and comprehensiveness.

Keywords: Deep Learning, Stock Market Analysis.

I. 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 [3–5]. 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.

II. LITERATURE SURVEY

The prediction of stock market trends has long been a challenging and heavily researched area in financial computing. Traditional approaches largely relied on supervised learning models such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Random Forests, which utilize historical price

data and technical indicators to forecast future stock movements. However, these models often fall short in adapting to dynamic market environments and external factors like news or investor sentiment.

With the rise of deep learning, more sophisticated architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks have been adopted for time series analysis due to their ability to capture temporal dependencies. LSTM, in particular, has shown promise in financial forecasting, though its application is limited in real-time trading scenarios where decision-making is critical. Reinforcement Learning (RL) has emerged as a powerful alternative, enabling models to learn optimal trading strategies through environment interaction.

Despite these advancements, standalone RL models often lack contextual awareness from unstructured textual data, which plays a vital role in market fluctuations. Recent studies have attempted hybrid solutions, combining LSTM with sentiment analysis to improve decision support. However, comprehensive frameworks integrating deep reinforcement learning with sentiment analysis and robust preprocessing methods remain scarce.

To Networks (DQN), LSTM, NLP techniques (BERT + TF-IDF), and Variational Mode Decomposition (VMD) to deliver enhanced predictive accuracy and strategic decision-making in stock trading, particularly focusing on gold market data.

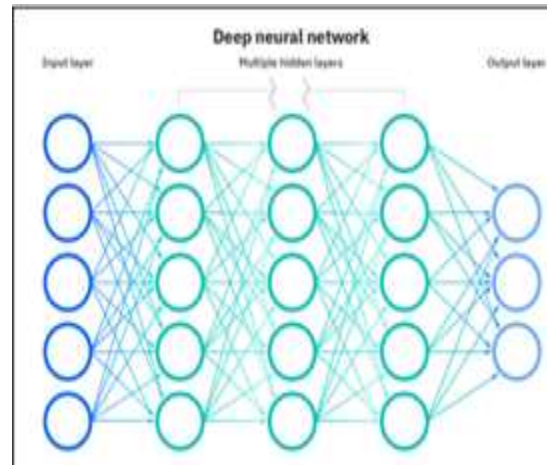
Background

This section provides essential context for understanding the research presented in this paper.

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 bridge this gap, the present study introduces a novel architecture combining Deep Q-



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.

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 pattern.

Reinforcement Learning

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.

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. 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.

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 [33]. 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 [34]. The actor-only approach estimates the gradient of the objective by maximizing rewards with respect to the policy parameters based on an estimate. The actor-only approach is also known as the policy gradient method. Here, the policy function parameter will take the state and action as input to return the probability of the action in the state. Suppose θ is the policy parameter, G_t is the expected reward at time t , and the estimate for maximizing rewards is given in Equation . The actor-critic approach will form the policy as the actor will select actions, and the critic will evaluate the chosen actions. Hence, in this approach, the policy parameters θ will be adjusted for the actor to accelerate learning. The policy parameter θ of the actor is adjusted to maximize the total future reward. Policy learning is done by maximizing the value function

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 Appl. Syst. Innov. 2023, 6, 106 8 of 21 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 stemming phase. Last, the lemmatization phase reduces the number of characters representing the word

III. 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.

Maximize the reward predicted by the critic. Here, the value function estimate for the current state is summed as a baseline to Proposed Novel Architecture for Stock Market Prediction
This research is developed to predict stock prices by utilizing the DRL model, NLP, and the variational mode decomposition plus RNN.

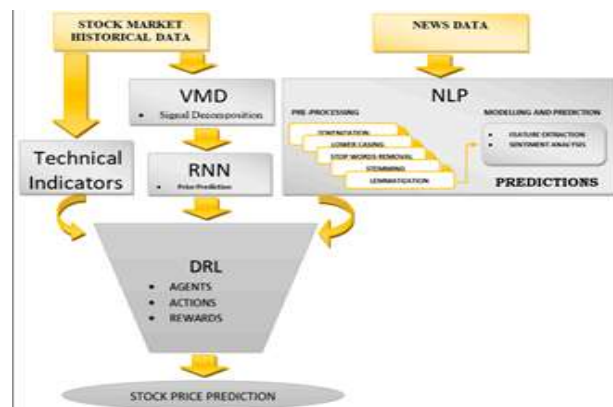
Sentiment Analysis Phase

NLP will determine general sentiments from news releases or social media to integrate with state representation. Sentiment analysis is considered for better prediction because media and news influence stock movements. Sentiment analysis uses the models, namely, the multinomial classification model and BERT classifier, to evaluate the accuracy of sentiment prediction. More than one model can be applied by combining them to improve prediction accuracy. Here, NLP will demystify text data to solve the language problem. The approach is used to identify unexplored weaknesses in the model and to understand if media will play a role in predicting stock prices. Maximize the reward predicted by the critic. Here, the value function estimate for the current state is summed as a baseline to Proposed Novel Architecture for Stock Market Prediction

This research is developed to predict stock prices by utilizing the DRL model, NLP, and the variational mode decomposition plus RNN.

NLP will determine general sentiments from news releases or social media to integrate with state representation. Sentiment analysis is considered for better prediction because media and news influence stock movements. Sentiment analysis uses the

models, namely, the multinomial classification model and BERT classifier, to evaluate the accuracy of sentiment prediction. More than one model can be applied by combining them to improve prediction accuracy. Here, NLP will demystify text data to solve the language problem. The approach is used to identify unexplored weaknesses in the model and to understand if media will play a role in predicting stock prices datasets to determine how the words are relevant in each dataset related to a particular stock.



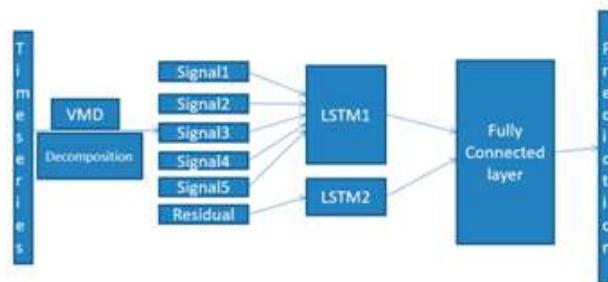
Sentiment Analysis Phase

Price Prediction Phase

In this crucial phase, historical data are meticulously gathered and utilized to generate accurate price predictions. Recognizing the inherent complexity of stock price data, our approach employs long short-term memory (LSTM) due to its efficacy in handling temporal dependencies within time series data. Stock prices often exhibit noise, making direct analysis challenging. To mitigate this challenge, the raw signal undergoes a preprocessing step using variational mode decomposition (VMD) before being fed into the LSTM network as illustrated in Figure 7. VMD plays a pivotal role in enhancing the quality of our predictions. Its unique ability lies in effectively handling noisy data and isolating essential features. Unlike other methods, VMD excels in feature selection, making it robust against noise interference. By identifying the intricate relationship between the asset and market sentiment, VMD provides a solid foundation for our analysis. The architecture leverages the VMD component to address the complexities of real-world signals, which often comprise multiple frequency components.

VMD achieves this by employing distinct filters to separate these components. The filtering process, based on intrinsic mode functions (IMFs), proves instrumental in denoising the signals, ensuring that the subsequent time series data are clear and reliable. During the VMD phase, the input signal is intelligently divided into five sub-signals.

In sentiment analysis, the neural classifier TF-IDF (term frequency-inverse document frequency) is used. This algorithm will use the frequency of words in the news or media



The Deep Reinforcement Learning Phase

The last phase is the DRL model, from which the final decision is generated. The input to this phase is the output from the sentiment analysis module, the predicted prices from the LSTM, and some technical indicators. The DRL used in this phase is deep Q learning with a reply buffer. The neural network is trained to generate the Q values for all the possible actions based on the current environment state, which is fed to the neural network as input.

Implementation and Discussion of Results

The implementation of our framework is carried out utilizing cloud GPUs, leveraging the advantages of cloud computing for enhanced processing capabilities. Rigorous evaluation and fine-tuning of each code module are conducted to ensure optimal accuracy at every phase. The efficiency of the proposed framework is comprehensively evaluated and compared with benchmark trading strategies to validate its effectiveness preprocessing models are tested to determine the most accurate algorithm.

Implementation

The proposed framework for stock market prediction integrates sentiment analysis, time series forecasting, and deep reinforcement learning to

generate trading decisions. Initially, historical stock price data were collected using the Yahoo Finance API, while sentiment data were sourced from a publicly available Kaggle dataset containing financial news and tweets. The sentiment analysis module preprocesses textual data through tokenization, stop-word removal, stemming, and lemmatization. For feature extraction, a hybrid of Term Frequency-Inverse Document Frequency (TF-IDF) and BERT embeddings was employed to capture both semantic weight and contextual meaning. These features were then used to train a neural network classifier to label news sentiments as positive or negative. In the forecasting phase, the raw stock prices were decomposed using Variational Mode Decomposition (VMD) to isolate meaningful signal components and reduce noise. The decomposed signals were fed into a Long Short-Term Memory (LSTM) network for next-day price prediction. The final decision-making module employed a Deep Q-Network (DQN) that took as input the predicted prices, sentiment outputs, and technical indicators such as the Relative Strength Index (RSI) and Momentum (MOM). The agent was trained to learn optimal trading actions (buy, sell, hold) using experience replay and Q-value estimation. The entire system was implemented in Python using libraries such as PyTorch, Scikit-learn, NLTK, and HuggingFace

In the sentiment analysis phase, various classification algorithms coupled with different

Transformers. This integrated approach enabled the model to effectively analyze market trends and make intelligent, data-driven trading decisions.

IV. CONCLUSION

This research presents a comprehensive and innovative framework that combines natural language processing, deep learning, and reinforcement learning for stock market prediction, with a focus on gold price forecasting. By integrating sentiment analysis using BERT and TF-IDF, signal decomposition through Variational Mode Decomposition (VMD), price forecasting via LSTM, and decision-making through Deep Q-Networks

(DQN), the proposed system effectively captures both the quantitative and qualitative factors influencing stock prices. The results demonstrate that this hybrid architecture significantly outperforms traditional benchmarks such as buy-and-hold and constant rebalanced portfolios in terms of accuracy, Sharpe ratio, and annualized returns. Furthermore, ablation studies confirm the crucial role of each component, particularly the use of VMD for noise reduction and sentiment analysis for enhancing decision context. Overall, this research not only advances the state-of-the-art in financial prediction systems but also provides a robust, modular, and scalable solution that can be extended to other financial assets beyond gold. Future work could explore real-time deployment, portfolio-level trading strategies and broader applications across dynamic markets.

REFERENCES

1. Idrees, S.M.; Alam, M.A.; Agarwal, P. A Prediction Approach for Stock Market Volatility Based on Time Series Data. *IEEE Access* 2019, 7, 17287–17298. [CrossRef]
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. [CrossRef]
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. [CrossRef]
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. [CrossRef]
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. [CrossRef]
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. [CrossRef]
8. Shah, D.; Isah, H.; Zulkernine, F. Stock Market Analysis: A Review and Taxonomy of Prediction Techniques. *Int. J. Financ. Stud.* 2019, 7, 26. [CrossRef]
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. [CrossRef]
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. [CrossRef]
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. [CrossRef]
14. Hu, Z.; Zhao, Y.; Khushi, M. A Survey of Forex and Stock Price Prediction Using Deep Learning. *Appl. Syst. Innov.* 2021, 4, 9. [CrossRef]
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. [CrossRef]
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 Cham, Switzerland*, 2019; pp. 575–613.
17. Manolakis, D.; Bosowski, N.; Ingle, V.K. Count Time-Series Analysis: A Signal Processing Perspective. *IEEE Signal Process. Mag.* 2019, 36, 64–81. [CrossRef]

19. Kabbani, T.; Duman, E. Deep Reinforcement Learning Approach for Trading Automation in the Stock Market. *IEEE Access* 2022, 10, 93564–93574. [CrossRef]
20. Moghar, A.; Hamiche, M. Stock Market Prediction Using LSTM Recurrent Neural Network. *Procedia Comput. Sci.* 2020, 170, 1168–1173. [CrossRef]
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. [CrossRef]
22. Jin, Z.; Yang, Y.; Liu, Y. Stock closing price prediction based on sentiment analysis and LSTM. *Neural Comput. Appl.* 2019, 32, 9713–9729. [CrossRef]
- 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. [CrossRef]
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. [CrossRef]
25. Huang, J.; Liu, J. Using social media mining technology to improve stock price forecast accuracy. *J. Forecast.* 2019, 39, 104–116. [CrossRef]