IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 11 April 2024, accepted 4 May 2024, date of publication 10 May 2024, date of current version 28 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3399548

RESEARCH ARTICLE

Stock Market Prediction With Transductive Long Short-Term Memory and Social Media Sentiment Analysis

ALI PEIVANDIZADEH[1], SIMA HATAMI[2], AMIRHOSSEIN NAKHJAVANI[3], LIDA KHOSHSIMA[4], MOHAMMAD REZA CHALAK QAZANI[5], MUHAMMAD HALEEM[6], AND ROOHALLAH ALIZADEHSANI[7], (Member, IEEE)

[1]University of Houston, Houston, TX 77204, USA
[2]Raja University, Qazvin 34145-1177, Iran
[3]Computer Engineering-Software at Islamic Azad University of Mashhad, Mashhad, Iran
[4]Department of Islamic Economics, Faculty of Social Sciences, Raja University, Qazvin, Iran
[5]Faculty of Computing and Information Technology, Sohar University, Sohar 311, Oman
[6]Department of Computer Science, Kardan University, Kabul, Afghanistan
[7]Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Waurn Ponds, VIC 3216, Australia

Corresponding author: Muhammad Haleem (m.haleem@kardan.edu.af)

ABSTRACT In an era dominated by digital communication, the vast amounts of data generated from social media and financial markets present unique opportunities and challenges for forecasting stock market prices. This paper proposes an innovative approach that harnesses the power of social media sentiment analysis combined with stock market data to predict stock prices, directly addressing the critical challenges in this domain. A major challenge in sentiment analysis is the uneven distribution of data across different sentiment categories. Traditional models struggle to accurately identify fewer common sentiments (minority class) due to the overwhelming presence of more common sentiments (majority class). To tackle this, we introduce the Off-policy Proximal Policy Optimization (PPO) algorithm, specifically designed to handle class imbalance by adjusting the reward mechanism in the training phase, thus favoring the correct classification of minority class instances. Another challenge is effectively integrating the temporal dynamics of stock prices with sentiment analysis results. Our solution is implementing a Transductive Long Short-Term Memory (TLSTM) model that incorporates sentiment analysis findings with historical stock data. This model excels at recognizing temporal patterns and gives precedence to data points that are temporally closer to the prediction point, enhancing the prediction accuracy. Ablation studies confirm the effectiveness of the Off-policy PPO and TLSTM components on the overall model performance. The proposed approach advances the field of financial analytics by providing a more nuanced understanding of market dynamics but also offers actionable insights for investors and policymakers seeking to navigate the complexities of the stock market with greater precision and confidence.

INDEX TERMS Stock market, sentiment analysis, unbalanced classification, proximal policy optimization, transductive long short-term memory.

I. INTRODUCTION

The stock market is a crucial component of a nation's economy, influencing its growth or decline through its performance over time [1]. Given the inherent unpredictability of financial markets, it is still being determined whether investments will yield profits or result in significant losses for investors. As a critical aspect of economic liberalization, stock markets play a vital role in the financial strategies of the global corporate world. Investors face the crucial decision of buying, selling, or holding a stock. Making the right investment choices can lead to substantial profits, but wrong

The associate editor coordinating the review of this manuscript and approving it for publication was Daniel Augusto Ribeiro Chaves.

decisions can lead to losses, affecting both the individual investor and the country's economy. Therefore, there is a pressing need for predictive models that can forecast stock prices with greater accuracy and efficiency.

A wealth of research has shown that including data from fundamental analyses, such as financial news on websites or social media postings, can improve the accuracy of models predicting stock prices [2], [3]. Moreover, between one-third and two-thirds of investors turn to social media platforms to collect information and gain insights into companies they are interested in. This reliance on social media not only shapes their investment strategies but also underscores the influence that online commentary can have on the fluctuations of stock prices. Integrating digital discourse into investment analysis represents a paradigm shift in gauging and understanding market sentiments. With their real-time updates and broad user engagement, social media platforms offer a dynamic and rich source of investor sentiment and market perception. This digital pulse can provide early signals of shifts in stock market trends, offering investors a competitive edge. As such, the ability to analyze and interpret social media sentiment is becoming an increasingly valuable tool in the financial analyst's toolkit, enabling more nuanced and informed decision-making in the fast-paced world of stock trading [4].

In stock market prediction, various time-series methodologies, including Long Short-Term Memory (LSTM)-based models, have been employed to develop predictive frameworks [5]. Although LSTM models have proven their worth across numerous sequence learning tasks, their global modeling approach, which relies on the entirety of the training data, may sometimes overlook subtle nuances within certain feature space areas. To address this limitation, the TLSTM model incorporates a transductive learning process, enhancing model sensitivity to minor variations in data [6]. By dynamically adjusting weights based on the proximity of data points to unknown values, TLSTM [7] offers a more nuanced and effective method for time-series prediction. This model melds the adaptive nature of transductive learning with LSTM's robustness, enabling it to adeptly navigate both local idiosyncrasies and overarching temporal patterns. Such an amalgamation of global and localized insights significantly bolsters the model's capacity to decipher and anticipate the intricacies of complicated time series, which require a delicate balance between recognizing immediate, specific patterns and understanding broader, temporal dynamics [8].

Few studies have ventured into using sentiment analysis for predicting stock market trends [9], [10], [11], but they often run into the problem of imbalanced data, which can compromise their effectiveness. To address this, strategies are put in place at both the data handling and algorithmic stages [12]. Data-level tactics involve adjusting the dataset to balance it out, such as decreasing the size of overrepresented groups, increasing the presence of underrepresented ones, or creating new data points to even the scales [13], [14]. Algorithmic-level strategies focus on fine-tuning the learning algorithms to better recognize and value the input from less represented groups, which might be infrequent but are often crucial for accurate predictions [15]. The main issue at the data level is the uneven distribution of data categories, leading to a bias towards more common outcomes and neglecting rarer, yet significant, market signals. At the algorithmic stage, the challenge lies in modifying the learning process to ensure that these vital but less common signals are not missed, enhancing the model's overall predictive power and reliability [16].

The rise of Deep Reinforcement Learning (DRL) has garnered considerable interest for its capability to handle complex tasks, particularly in scenarios where classes are imbalanced [17], [18]. DRL boosts the effectiveness of classification systems by reducing noise and enhancing relevant features, showing success in various fields. One of the critical strengths of DRL is its ability to adapt its learning approach based on rewards, which is especially useful for addressing data imbalances. By tailoring the reward system to prioritize accurately identifying less represented classes, DRL models can focus more on these crucial yet often overlooked samples. This strategy helps maintain the model's impartiality towards dominant classes, enhancing detection accuracy. However, DRL models can encounter challenges related to the bias-variance dilemma and their sensitivity to hyperparameters [19]. PPO [20], an advanced on-policy RL algorithm, overcomes some challenges. It uses a clipping mechanism to ensure stable training by moderating policy updates and preventing significant deviations from the current policy. PPO is known for its computational efficiency and is well-adapted for large-scale or complex scenarios involving continuous variables. Various modifications like Trust Region-Guided PPO and Truly PPO [21] have been developed to boost PPO's efficiency. However, they often overlook the potential benefits of using off-policy data to enhance sample efficiency. Off-policy PPO has demonstrated remarkable achievements in gaming [22], robotics [23], and continuous control tasks [24] by optimizing policies with data gathered from agent interactions, offering greater sample efficiency than on-policy techniques. Utilizing off-policy data allows these methods to minimize the costs associated with extensive direct interactions, making them ideal for tackling complicated sequential decision-making problems in real-world settings. Off-policy approaches, such as Q-learning, store past experiences in a replay buffer, enabling the algorithm to learn from a broad range of past interactions. This increases the system's resilience and flexibility in changing environments and supports exploring various strategies, proving beneficial for applications that require long-term planning like autonomous navigation, financial decision-making, and complex industrial operations [25].

This study addresses two primary challenges in social media analysis for stock price prediction: the issue of unbalanced classification in sentiment analysis and the complexity of capturing temporal subtleties in financial time-series data. To tackle these challenges, we propose a novel two-stage methodology. In the first stage, we develop a sentiment

analysis model that employs three dilated convolution layers to extract feature vectors for classification. To overcome the problem of unbalanced classification, where particular sentiments or outcomes are disproportionately represented, we implement an innovative approach using Off-Policy PPO. This approach conceptualizes sentiment classification as a series of decision-making processes where an agent is rewarded for accurate classifications. To address the imbalance, we strategically assign lower rewards to the majority class than the minority class, encouraging the model to pay more attention to underrepresented categories. In the second stage, we utilize a TLSTM technique that combines sentiment analysis results with historical stock data for prediction. The TLSTM method excels at recognizing the intricate temporal patterns inherent in stock market data. By adopting transductive learning principles, the model allows samples closer to the test data to have a more pronounced impact on the prediction process, enhancing the accuracy of forecasting future stock prices.

The pivotal contributions of the proposed model are encapsulated in several key areas:

- Our novel model for forecasting stock market trends employs the power of TLSTM. This approach is distinct in its ability to gauge the influence of training examples on the cost function, a measurement made according to their similarity to the test point. The result is a considerable enhancement in the precision and efficacy of stock market predictions.
- We implement an original Off-Policy PPO-based technique to counterbalance the skewed classification challenges observed in sentiment analysis. This involves suggesting an adapted surrogate goal that uses off-policy data to prevent considerable updates to the policy.
- Our model integrates a unique, rewarding system where correct decisions are positively reinforced, and incorrect ones are penalized. We tackle the issue of data skewness by allotting increased incentives to the less represented category, thus motivating the model to allocate greater focus to frequently overlooked entries. This deliberate tactic assists in accomplishing a more equitable and unbiased categorization.

The rest of the article unfolds as follows: Section II presents a comprehensive review of the existing literature on stock market prediction. Section III delves deeper into the proposed methodology. Experimental results and their corresponding analyses are showcased in Section IV. Lastly, Section V wraps up the article with the conclusion.

II. RELATED WORK

The field of stock market prediction has been rigorously explored, with numerous surveys providing comprehensive reviews [26], [27], [28]. These surveys offer diverse perspectives, ranging from the efficacy of classical forecasting techniques to applying advanced machine learning models like LSTM networks and integrating sentiment analysis for enhanced accuracy. In this section, we delve into the existing literature on stock market forecasting, categorizing it into three main areas: traditional forecasting methods, LSTM-based models, and the role of sentiment analysis in predicting market trends.

A. CONVENTIONAL TIME-SERIES ANALYSIS

Some efforts have been made to forecast stock market trends using classic time-series analysis. Sangeetha and Alfia [29] suggested applying Machine Learning with Evaluated Linear Regression (ELR-ML) to predict the S&P 500 index, utilizing factors like opening, closing, low, high, and volume. This method takes advantage of machine learning's ability to navigate the complex and unpredictable nature of the stock market, marking a significant step forward in making stock market predictions more automated and data-driven. Zhang et al. [30] introduced an innovative technique to boost the accuracy of stock market forecasts by combining a wavelet soft-threshold de-noising model with Support Vector Machine (SVM) classification. This strategy improved the clarity of the training data by removing stochastic trend noise from the Shanghai Stock Exchange (SSE) Composite Index, achieving a 60.12% success rate in predictions, a notable improvement over the 54.25% accuracy obtained with noisy data. Mahmoodi et al. [31] developed a strategy to improve the prediction of stock market trading signals, considering the market's inherent volatility and unpredictability. They employed an SVM enhanced with Particle Swarm Optimization (PSO) for more effective and precise categorization. Kurani et al. [32] examined how financial organizations depend on computational technology for various activities, from managing budgets to forecasting stock trends. They explored the roles of Artificial Neural Networks (ANNs) and SVM in making predictions, emphasizing the robustness of ANN to incomplete data and the efficiency of SVM in avoiding overfitting by using straightforward decision boundaries.

Traditional time-series methods assume a linear connection between stock prices, working best with stable and predictable trends. However, these methods need to be revised regarding the stock market's more complex, nonlinear relationships. Furthermore, the stock market is influenced by numerous complex factors that simple time-series analyses need to pay more attention to, leading to less effective predictions.

B. LSTM MODEL

Numerous studies have shown that LSTM networks can improve time-series forecasting [33], [34], [35]. Behura et al. [36] introduced a cutting-edge method to predict stock prices using an advanced Multi-Layer Sequential LSTM model enhanced by the Adam optimizer. This technique uses normalized data broken down into time intervals to link past and future values while overcoming common issues like the vanishing gradient problem in essential neural networks. Koo and Kim [37] developed the Centralized Clusters

Distribution (CCD) as a new technique to filter input data, significantly improving Bitcoin price predictions by addressing the price's extreme variability. Combined with the Weighted Empirical Stretching (WES) loss function, which adjusts penalties based on data distribution, this method significantly boosts prediction accuracy. Bilhah et al. [38] conducted a thorough analysis of LSTM networks for forecasting in the unpredictable stock market, utilizing a broad array of real data to showcase LSTM's superiority over traditional methods. They aimed to provide a basis for comparing asset values by analyzing historical market data. Chen et al. [39] focused on data preprocessing for stock price predictions, introducing 57 technical indicators to capture economic signals better and employing the Least Absolute Shrinkage and Selection Operator (LASSO) algorithm to select the most relevant features. Their use of Ca-LSTM (cascade LSTM) aimed to extract more nuanced information from the data to improve prediction reliability. Sivadasan et al. [40] applied recurrent neural networks, including Gated Recurrent Unit (GRU) and LSTM, to discern patterns in stock market data, using a sliding window method to examine daily stock metrics and integrate various technical indicators to refine their models further. Chadidjah et al. [41] reviewed the efficacy of LSTM networks in forecasting Apple Inc.'s stock prices by comparing different LSTM configurations. This study assessed how well these models could detect stock data trends, taking into account their computational efficiency and predictive accuracy.

Traditional LSTM methods have made significant strides in stock market forecasting, but they frequently must focus on subtle variances within specific areas of the feature space. This is due to the development of a comprehensive model encompassing the entire training dataset, which could restrict their usefulness in real-world stock market scenarios. Our research introduces a new strategy that leverages the advanced features of TLSTM to mitigate these drawbacks. This combined approach improves upon the standard LSTM models, providing a more powerful and flexible tool for stock market analysis.

C. SENTIMENT ANALYSIS

Sentiment analysis fundamentally explores and quantifies the emotional tone behind a body of text. This approach has been widely applied in various fields to understand better the attitudes, opinions, and emotions expressed in written language [42]. In the context of stock markets, sentiment analysis is increasingly recognized for its capacity to unveil market participants' underlying moods and sentiments, potentially influencing stock prices and market trends [43]. Drawing from the intersection of behavioral finance and data science, sentiment analysis examines how collective emotions and investor sentiments, often disseminated through social media and financial news, can predict market movements [44], [45].

While some studies have established a link between the sentiment of online commentary and stock trends, few have ventured into predicting specific stock prices through sentiment analysis [46], [47]. Bassant et al. [48] proposed a unique approach to improve stock market movement predictions by refining sentiment analysis with neutrosophic logic (NL), which adeptly manages uncertain and indeterminate data. This method involves classifying social media sentiments more accurately and using these insights and historical stock data as inputs for a deep-learning LSTM model to forecast stock movements over a set period. BL and BR [49] introduced a pioneering framework for predicting stock prices, combining market sentiment data and news sources. This method employs technical stock indicators like Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), and Moving Average (MA) and applies sentiment analysis on news content using techniques such as keyword extraction, sentiment categorization, and holo-entropy-based feature extraction, all processed through a deep neural network. The precision of this model is further refined by a self-improved Whale Optimization Algorithm (SIWOA) that trains the neural network and a Deep Belief Network (DBN) that makes the final stock predictions, with SIWOA adjusting the DBN's parameters. Swathi et al. [9] developed an innovative sentiment analysis strategy for forecasting stock prices using Twitter data. This approach merges a Teaching and Learning Based Optimization (TLBO) model with LSTM networks. It processes tweets to determine their sentiment towards stock prices. It employs the Adam optimizer to fine-tune the LSTM's learning rate, with the TLBO model enhancing the LSTM's output for more accurate stock price predictions based on social media sentiments. Wu et al. [50] devised a stock price prediction technique that incorporates a variety of data sources, including historical stock information, technical indicators, and unconventional sources such as financial news and stock forums. They used convolutional neural networks for sentiment analysis to assess investor sentiment. They combined this with LSTM networks to achieve higher prediction accuracy in the China Shanghai A-share market. Harguem et al. [11] explored the influence of Twitter sentiment on global corporate stock prices by analyzing tweets for their positive, negative, and neutral tones. Using data from the NASDAQ 100, they optimized a subset of data and applied One-Hot Encoding for feature simplification. They modeled the correlations with SVM algorithms using various kernels and implemented cross-validation to verify the model's precision and reliability, achieving notable accuracy with the Linear kernel. Ye et al. [51] proposed a cutting-edge ensemble deep learning model for predicting Bitcoin prices within the next 30 minutes, utilizing price data, technical indicators, and sentiment indices. This model integrates LSTM and GRU neural networks with a stacking ensemble technique to improve accuracy. The model's sentiment analysis component uses social media text analyzed through linguistic and statistical methods and technical indicators for a

thorough forecasting methodology. Gupta et al. [52] proposed a machine learning-based method to improve investment decisions in the stock market by accurately forecasting future stock prices. They recommended using historical data and sentiment analysis from news articles and employing LSTM networks. This strategy recognizes the strong relationship between stock price movements and news coverage, aiming to offer investors more dependable advice on whether to buy, sell, or hold stocks.

Existing methods frequently need help with the issue of imbalanced class distribution. This study introduces our sentiment analysis model, which employs off-policy PPO to handle imbalanced classification challenges effectively.

III. MATERIALS AND METHODS

Our model, crafted to predict stock market prices, progresses through a structured four-step methodology. It starts with gathering extensive data from social media and financial markets, which is then meticulously preprocessed to organize and refine for further analysis. The next stage involves conducting an in-depth semantic analysis to assess public sentiment derived from social media platforms. The culmination of this process is the prediction of stock market trends, achieved by integrating these sentiment insights with historical stock data.

We have chosen the Off-policy PPO algorithm to address a significant issue in sentiment analysis: the prevalent imbalance between more common (majority) and less common (minority) classes during data classification. Off-policy PPO effectively counters this imbalance by altering the reward system in the training phase, thus promoting the precise classification of lesser-represented sentiments and ensuring a fairer analytical approach.

To overcome the challenge of merging sentiment analysis outcomes with the dynamic nature of stock prices, we utilize the TLSTM model. The TLSTM model is particularly effective at identifying and giving importance to temporal patterns, focusing on data points closer to the target prediction timeframe. This capability significantly enhances the precision of our stock market predictions.

A. DATA COLLECTION

Our research utilizes an extensive dataset that combines financial news and stock market figures from January 2015 to December 2020. This dataset contains around 12000 daily news articles from leading financial news outlets like Moneycontrol, India Infoline Finance Limited (IIFL), and Economic Times and from social media, specifically Twitter [53]. We chose these sources for their authoritative and dependable coverage of market-relevant information, including company updates, industry trends, and economic news, which are crucial for analyzing stock market dynamics.

In addition, we collected stock market information from the National Stock Exchange (NSE) of India, focusing on a carefully chosen group of 50 stocks and 10 Exchange-Traded Funds (ETFs) covering a wide range of industries. Our predictive analysis is conducted on individual stock tickers within this group rather than broader index funds. This selection was made based on factors like market capitalization and liquidity to represent the Indian market's diversity accurately. The stock data includes daily details of opening, high, low, and closing prices and trading volume, totaling over 1.1 million entries.

An analysis of the sentiment in news articles showed that neutral tones were most common (60%), with positive (25%) and negative (15%) sentiments following. This pattern reflects the typically cautious optimism seen in financial reporting. Our correlation study between news sentiment and stock price trends found a slight positive relationship in stock price gains, especially in the tech and pharmaceutical sectors, when news sentiment was positive. Additionally, our examination of trading volumes identified significant increases aligned with major company news or policy adjustments, demonstrating the market's responsiveness to news events.

B. DATA PREPROCESSING

The pre-processing phase began with the synchronization of the text, converting it into lowercase characters. It is essential to consider various factors present in the text during this stage, as they can significantly impact the classification process. The primary objective of this pre- processing is to prepare the input data for the sentiment classifier, ensuring it is in a suitable format for analysis. This pre-processing procedure encompasses several tasks, including removing links, special symbols, and emoticons, as well as eliminating stop words. Additionally, it involves analyzing the parts of speech, applying stemming techniques and conducting tokenization to break the text into individual units. These steps collectively contribute to refining the data and extracting meaningful features for sentiment analysis. The subsequent sentiment classifier method relies on the output generated by the preprocessor subsystem.

The methodology described in this document leveraged a data-centric strategy by harnessing opening and closing stock price records. This extensive dataset laid the groundwork for exploring and comprehending stock price trends over periods. The methodology adopted a sentiment polarity technique to depict these stock prices effectively. Its goal was to encapsulate the emotional sentiment tied to the stock market by classifying the prices into positive and negative categories. This sentiment polarity framework offered crucial perspectives on market behavior, facilitating a deeper grasp of how feelings and perceptions impact stock prices. Research was undertaken at Stanford University to ascertain the efficacy of this strategy [54]. The research experiments incorporated the sentiment polarity figures obtained from the methodology. These tests examined the link between sentiment polarity and stock price fluctuations, illuminating the connection between market mood and economic results. The research outcomes underscored the importance of sentiment evaluation in forecasting stock market directions and yielded insightful implications for investors, market participants, and economic researchers.

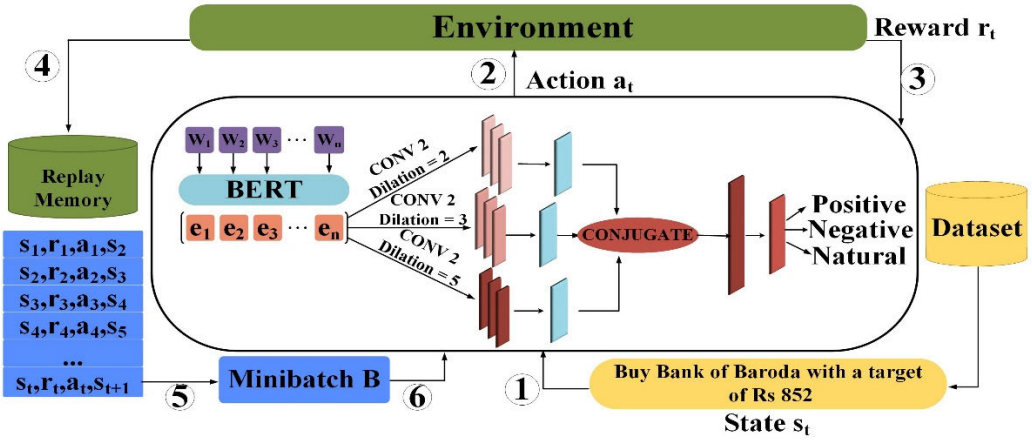


FIGURE 1. Overview of the proposed semantic analysis model. In phase 1, an instance s_t is initially extracted from the collection and fed into the framework. Subsequently, in phase 2, the operation a_t is relayed back to the milieu to obtain the next instance s_{t+1} along with the incentive r_t in phase 3. Phase 4 archives the sequence of $\{s_t, a_t, r_t, s_{t+1}\}$ within the replay archive. After the replay archive has gathered a multitude of sequences, phase 5 involves the selection of a stochastic minibatch of sequences to refine the framework parameters in phase 6. This cycle is perpetuated until the framework precisely categorizes the input instances.

C. SEMANTIC ANALYSIS

Figure 1 illustrates the structure of the proposed semantic analysis model. The model accepts a sentence $D = [w_1, w_2, \ldots, w_n]$ as its input, with w_i denoting the words and n representing the maximum count of words a sentence can have. This input is subsequently fed into the Bidirectional Encoder Representations from Transformers (BERT) model. The result from the BERT model is an embedding matrix $M = [e_1; e_2; \ldots; e_n]$, where each e_i is the embedded representation of the corresponding word w_i. The matrix M is subjected to three parallel dilated convolution layers to obtain features from the sentence. Each of these branches independently derives a feature vector from the sentence. Post convolution, max pooling is employed to isolate the most significant features and reduce the computational load of the network. The output derived from the max pooling layers is subsequently channeled into the Multilayer Perceptron (MLP) network for classification. The output generated by the MLP is a vector of length three since every sentence s must be categorized into one of three classes: positive, negative, or neutral. However, since most sentences are often classified as positive, the classifier encounters imbalanced classification, decreasing system performance. The Off-Policy PPO algorithm is employed to address this issue, creating a sequential decision-maker to overcome the imbalanced classification problem.

1) OFF-POLICY PPO

To address the challenge of inefficiently using samples within the PPO technique, we present an Off-Policy PPO approach that integrates off-policy data to optimize policies. Inefficient sample utilization is a notable hurdle in reinforcement learning, where agents often require numerous interactions with the environment to acquire effective policies. This issue becomes particularly prominent when facing real-world situations or resource-intensive simulations. Although effective the conventional on-policy PPO method exhibits considerable sample inefficiency, as it updates policies exclusively based on current data collected during interactions with the environment. Conversely, our Off-Policy PPO strategy tackles this drawback by capitalizing on off-policy data, encompassing information gathered from prior interactions with the environment or alternative policies. Using such data, the agent can capitalize on past experiences more effectively, diminishing the necessity for redundant exploration and augmenting overall learning efficiency. The primary challenge in integrating off-policy data is maintaining stability during policy updates. Given the diversity of off-policy data, directly employing it for policy optimization might lead to unstable learning, resulting in subpar convergence or even divergence.

To mitigate these concerns, we introduce a clipped surrogate objective that balances the trade-off between exploration and exploitation. This clipped surrogate objective empowers us to harness off-policy data while limiting policy updates to a reasonable range. This restriction prevents abrupt policy changes that could destabilize the learning process. Furthermore, our Off-Policy PPO approach retains the advantageous attributes of the original PPO algorithm, such as guarantees of monotonic improvement and straightforward implementation. However, it significantly amplifies learning efficiency, rendering it more appropriate for real-world applications where data collection could be expensive, time-consuming, or associated with risks. The maximization challenge that can leverage off-policy data in the Off-Policy Trust Region Policy Optimization (TRPO) [55] is:

$$\max_{\pi} E_{s\sim\rho_\mu, a\in\mu}\left[\frac{\pi(a\mid s)}{\mu(a\mid s)} A_{\pi_{\theta_i}}(s,a)\right] \quad (1)$$

$$st.\ \bar{D}_{KL}^{\rho_\mu, sqrt}(\mu, \pi_{\theta_i})\bar{D}_{KL}^{\rho_\mu, sqrt}(\pi_{\theta_i}, \pi) + \bar{D}_{KL}^{\rho_\mu}(\pi_{\theta_i}, \pi) \le \delta \quad (2)$$

$$\rho_{\mu}(s) = \sum_{t=0}^{\infty} \gamma^{t} P(s_t = s | s_0, \mu) \tag{3}$$

$$\bar{D}_{KL}^{\rho_{\mu}}\left(\pi_{\theta_i}, \pi\right) = E_{s\sim\rho_{\mu}}[D_{KL}\left(\pi_{\theta_i}(.\,|\,s)\,||\pi(.\,|\,s)\right)] \tag{4}$$

$$\bar{D}_{KL}^{\rho_{\mu},sqrt}\left(\mu, \pi_{\theta_i}\right) = E_{s\sim\rho_{\mu}}[\sqrt{D_{KL}\left(\mu(.\,|\,s)\,||\pi_{\theta_i}(.\,|\,s)\right)}] \tag{5}$$

$$\bar{D}_{KL}^{\rho_{\mu},sqrt}\left(\pi_{\theta_i}, \pi\right) = E_{s\sim\rho_{\mu}}[\sqrt{D_{KL}\left(\pi_{\theta_i}(.\,|\,s)\,||\pi(.\,|\,s)\right)}] \tag{6}$$

In the context where μ is defined as the behavior policy, the optimization problem aims to maximize the expected advantage of choosing actions from policy π over μ in states sampled according to the state distribution ρ_{μ}. Here, $\pi(a|s)$ and $\mu(a|s)$ represent the probability of taking action a in state s under policies π and μ, respectively, while $A_{\pi_{\theta_i}}(s,a)$ denotes the advantage function under the candidate policy parameterized by θ_i, indicating the relative benefit of action a in state s. The optimization is subject to a constraint on the sum of the square root of the average Kullback-Leibler (KL) divergences between μ and π_{θ_i} and between π_{θ_i} and π, along with the average KL divergence between π_{θ_i} and π, all weighted by the state distribution under μ, and bounded by a threshold δ. The state distribution $\rho_{\mu(s)}$ is defined as the discounted sum of probabilities of reaching state s from an initial state s_0 under policy μ, with γ as the discount factor. The average KL divergence terms, $\bar{D}_{KL}^{\rho_{\mu}}\left(\pi_{\theta_i}, \pi\right)$ and its square root versions quantify the expected divergence in policy behavior in states sampled according to ρ_{μ}, capturing the degree of deviation between the policies across the state space.

Without the constraint outlined in Equation 2, maximizing the mentioned surrogate objective using off-policy data in Equation 1 leads to an overly significant alteration in the policy. To address this issue, a straightforward solution involves adopting the clipping method from PPO to adjust the surrogate objective in Equation 3:

$$L_{\mu}(\pi) = E_{s\sim\rho_{\mu},a\in\mu}[\frac{\pi(a\,|\,s)}{\mu(a\,|\,s)} A_{\pi_{\theta_i}}(s,a)] \tag{7}$$

Having $L_{\mu}(\pi)$ as defined in Equation 7, the related truncated surrogate goal utilizing off-policy data becomes:

$$\overline{L_{\mu}}(\pi) = E_{s\sim\rho_{\mu},a\in\mu}[min(\frac{\pi(a\,|\,s)}{\mu(a\,|\,s)} A_{\pi_{\theta_i}}(s,a)\,, clip(\frac{\pi(a\,|\,s)}{\mu(a\,|\,s)}, 1-\in, 1+\in) A_{\pi_{\theta_i}}(s,a))] \tag{8}$$

Notably, the proportion of the policy $\frac{\pi(a\,|\,s)}{\mu(a\,|\,s)}$ commonly tends to be lower than $1-\in$ or higher than $1+\in$. Consequently, the desired policy $\pi(a\,|\,s)$ often remains unchanged and experiences no modifications during the optimization procedure of the truncated surrogate objective in Equation 8. To address this issue, we present an alternative truncated surrogate objective that adapts the lower and upper limits $((1-\in), (1+\in))$ in Equation 9 using a factor of $\frac{\pi_{\theta_i}(a\,|\,s)}{\mu(a\,|\,s)}$:

$$L_{Off-PolicyPPO}^{CLIP}(\pi)$$

$$= E_{s\sim\rho_{\mu},a\in\mu}[\min[\frac{\pi(a\,|\,s)}{\mu(a\,|\,s)} A_{\pi_{\theta_i}}(s,a)\,, clip\left(\frac{\pi(a\,|\,s)}{\mu(a\,|\,s)}, \frac{\pi_{\theta_i}(a\,|\,s)}{\mu(a\,|\,s)}(1-\in), \frac{\pi_{\theta_i}(a\,|\,s)}{\mu(a\,|\,s)}(1+\in)\right) \times A_{\pi_{\theta_i}}(s,a)\,]] \tag{9}$$

2) SETTING

In this article, we focus on implementing the Off-policy PPO algorithm within the context of sentiment analysis.

The subsequent description outlines how the methodology operates and provides an understanding of each element:

- State $\boldsymbol{s_t}$: This corresponds to the sample observed at time step t.
- Action $\boldsymbol{a_t}$: The categorization executed on the sample is regarded as an action. This signifies a choice carried out by the network, grounded in its prevailing comprehension of the objective.
- Reward $\boldsymbol{r_t}$: A reward is furnished for every categorization, and designed to steer the network towards accurate categorization. The formulation of this remuneration process is expressed as:

$$r_t(s_t, a_t, y_t) = \begin{cases} +1, a_t = y_t \text{ and } s_t \in D_O \\ -1, a_t \neq y_t \text{ and } s_t \in D_O \\ \lambda, a_t = y_t \text{ and } s_t \in D_N \\ -\lambda, a_t \neq y_t \text{ and } s_t \in D_N \end{cases} \tag{10}$$

where D_O and D_N denote the minority and majority classes correspondingly. Properly/mistakenly classifying a sample from the majority class leads to a positive/negative gain of $+\lambda$/ - λ. The outlined approach guides the network to prioritize accurately classifying instances of the scarcer class, assigning a higher absolute reward value. Simultaneously, including the majority class and the flexible reward parameter within the range of $0 < \lambda < 1$ brings complexity to the reward framework, allowing precise adjustment of the network's focus between the more frequent and less frequent classes.

We create the simulation environment according to the specified criteria. The design of the policy network has been thoughtfully formulated, considering both the intricacy and abundance of training instances. In this specific scenario, we meticulously structure the network input to match the training sample format, while the number of classes present in the instance data intelligently determines the output. Our model, outlined in Algorithm 1 (inspired by [56]), incorporates a comprehensive instructional approach that enables the agent to consistently engage in the learning process until it attains an optimal policy. The decision-making mechanism for action selection is guided by a self-interested policy, ensuring that the agent's decisions are influenced by its best interests. These actions are subsequently assessed using Equation 10, allowing us to gauge their efficacy quantitatively. To guarantee the robustness and effectiveness of our methodology, we iterate through the process for E iterations, which is set to 1500 in this article. After each iteration, we retain the policy network parameters, offering valuable

Algorithm 1 Pseudo-code of training the proposed semantic analysis model

Input: Training dataset $D = \{(s_1, y_1), (s_2, y_2), \ldots, (s_T, y_T)\}$, Learning rate α
Output: Updated policy parameters θ
$\theta, \theta_{old} \leftarrow$ Initialize policy parameters
$\phi \leftarrow$ Initialize value function parameters
$A \leftarrow$ Initialize advantage estimator
B $\leftarrow$ Initialize replay buffer
for e =1 to E **do**:
  Randomize the order of dataset D
  Set initial state to s_1
  for t =1 to T **do**:
    $a_t \leftarrow$ Choose action according to policy π_θ given state s_t
    $r_t \leftarrow$ Calculate reward using Reward (x_t, a_t, y_t)
    $s_{t+1} \leftarrow$ Determine the next state from dataset D
    B $\leftarrow$ Store transition $(s_t, a_t, r_t, s_{t+1}, \mu(\cdot|s_k))$
  end for
  for i =1 to N **do**:
    $(s_k, a_k, r_k, s_{k+1}, \mu(\cdot|s_k)) \leftarrow$ Draw random mini-batch from B
    $A(s_k, a_k) \leftarrow$ Estimate advantages using the value function with parameters ϕ
    $\theta \leftarrow$ Optimize policy π with objective $L^{CLIP}_{Off-PolicyPPO}(\pi)$ using α
    $\theta_{old} \leftarrow \theta$
  end for
end for

insights into the learning progress of the model and evolutionary trajectory. By meticulously adhering to this iterative approach, we can comprehensively evaluate the model's performance and make well-informed choices for further enhancement and refinement.

D. STOCK MARKET PREDICTION

LSTM networks are highly valued in stock market analysis for their ability to understand temporal sequences in data, making them ideal for predicting market prices and trends. Their unique architecture allows them to remember long-term dependencies, identify intricate patterns, and manage irregularities in data.

LSTM architecture [57] functions through a gating mechanism. This mechanism manages data retention over intervals, oversees the length of data retention, and determines the appropriate moments for data retrieval from the memory cell. For efficient data manipulation, LSTM utilizes three distinct gates (Refer to Figure 2). Equations 11-15 are applied to compute various elements: i_t (input gate), f_t (forget gate), o_t (output gate), c_t (memory cell), and h_t (hidden state) of the LSTM at a specific moment t, given the input x_t [58].

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \tag{11}$$
$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \tag{12}$$
$$c_t = f_t c_{t-1} + i_t tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{13}$$
$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \tag{14}$$
$$h_t = o_t tanh(c_t) \tag{15}$$

where $\sigma(.)$ represents the sigmoid function, acting as the gating mechanism's activation function. The full weight matrices, denoted as W_{xk} for $k \in \{i, f, o, c\}$, correspond to the weights linked to the input x_t across the input, forget, and output gates, as well as the memory cell. The weight matrices W_{ck} for $k \in \{i, f, o\}$ are formed as diagonal matrices, facilitating the connection between the memory cell and the various gates. It is important to highlight that the neuron count for all gates is set in advance, and equations 11 to 15 are executed for each neuron separately. Assuming n is the number of neurons, the sets $\{i_t, f_t, c_t, o_t, h_t\}$ are within the space $R^{n\times 1}$. For ease of discussion, we show the biases and weights in LSTM as b_{lstm} and w_{lstm}. The LSTM's operations are succinctly expressed as follows:

$$\begin{cases} c_t = f(c_{t-1}, h_{t-1}, x_t; w_{lstm}, b_{lstm}) \\ h_t = g(h_{t-1}, c_{t-1}, x_t; w_{lstm}, b_{lstm} \end{cases} \tag{16}$$

We now introduce the Transductive LSTM. Assuming z(η) as an unobserved sequence, the TLSTM state space formulation is expressed thus:

$$\begin{cases} c_{t,\eta} = f(c_{t-1,\eta}, h_{t-1,\eta}, x_t; w_{lstm,\eta}, b_{lstm,\eta}) \\ h_{t,\eta} = g(h_{t-1,\eta}, c_{t-1,\eta}, x_t; w_{lstm,\eta}, b_{lstm,\eta} \end{cases} \tag{17}$$

The structural blueprint delineated in Equation 17 significantly deviates from the one depicted in Equation 16. The parameters within Equation 16 stay unaffected by the assessment point; conversely, in Equation 17, these parameters are modulated by the feature vector of the assessment point. The notation η is utilized as a subscript to highlight the adaptability of the model's parameters upon including a novel data point, represented as z(η). It is crucial to underscore that the assessment label is considered unknown, and the sole role of the assessment point within the training phase is to gauge the relevance of the training data points based on the correlation between their feature.

In the present study, we employ a 15-day time frame for forecasting stock values, integrating various input features such as Open, High, Low, Close, Volume, and a semantic class to ensure a comprehensive analytical approach. These selected features, crucial indicators of stock price movement, and market sentiment serve as input variables for our TLSTM model. The TLSTM layers, adept at accommodating both global and local temporal dependencies and patterns, play a pivotal role in accurately predicting future stock prices, particularly in the context of this research.

It is imperative to note that the considered 15-day period represents a strategic selection to ensure that the model captures pertinent short-term fluctuations in stock values while accommodating the inherent noise and volatility in financial markets. Throughout this period, the model has been tasked with identifying and learning from the intrinsic patterns and tendencies within the market data, thereby refining its predictive capabilities.

The result produced by the TLSTM layers yields a single value, representing the forecast of the network for the market value of the following day. This predictive process is essential in rendering actionable insights for traders and financial

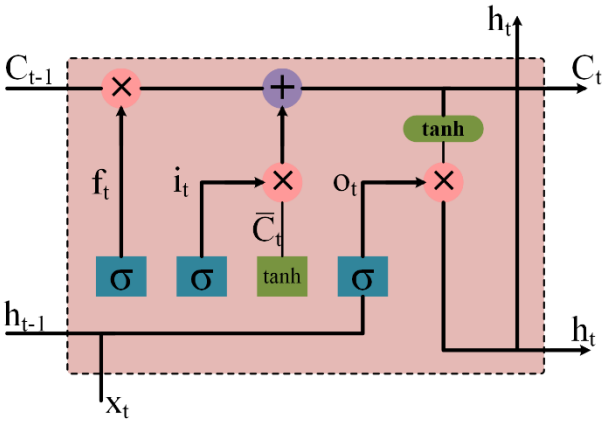


FIGURE 2. Structure of an LSTM cell.

analysts, enabling them to make well-informed decisions by gauging potential market movements. The semantic class, determined through a meticulous analysis of relevant textual data, further enriches the model by incorporating the impacts of market sentiment and related factors.

IV. EMPIRICAL EVALUATION

We employed the well-established cross-validation method to identify the optimal value for each hyperparameter, utilizing a 5-fold cross-validation. This strategy not only facilitates a thorough exploration of a wide array of hyperparameter configurations but also systematically examines the efficacy of each combination. Importantly, we ensured that while testing one hyperparameter, the rest were held constant to isolate the effects of individual parameter variations. Within each cross-validation iteration, the proficiency of the model is gauged against predetermined benchmarks, maintaining a uniform standard for performance evaluation. Following numerous iterations across different folds, this process grants an in-depth understanding of the performance landscape of the model across a diversity of configurations. This exhaustive analytical procedure enables a comparative assessment of all outcomes, thereby allowing us to pinpoint the most proficient set of hyperparameters demonstrating superior performance throughout the cross-validation phases. Leveraging this rigorous approach is essential in optimizing our model to achieve peak performance and ensuring its resilience and flexibility across varied scenarios. Figure 3 illustrates the critical hyperparameters incorporated in our proposed model, delineating different feasible values. Table 1 encapsulates the pivotal hyperparameters associated with the proposed models, offering a detailed enumeration of the potential value range for each.

The proposed stock market prediction model was compared against two conventional methods, ELR-ML [29], WD-SVM [30], two LSTM-based methods, MLS-LSTM [36], LSTM [38], and four sentiment analysis-based models, NL-LSTM [48], DL-SIWOA [49], GRU-LSTM [51], HiSA-SMFM [52]. Additionally, we juxtaposed the model against two derivative models: Proposed without Off-policy PPO, which doesn't use Off-policy PPO for classification, and Proposed with LSTM, substituting TLSTM with LSTM. In total, eleven distinct models were run for the empirical evaluation of our proposed methodology. Table 2 presents a summary of these comparative results. All models were assessed on a common dataset using metrics Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). It is important to highlight that the data preprocessing described in Section 8 has been uniformly applied across all methodologies.

The conventional methods, ELR-ML and WD-SVM, exhibit higher RMSE, MAPE, and MAE values than the LSTM-based and sentiment analysis-based models, indicating a relatively lower performance in stock market prediction. Specifically, WD-SVM shows a modest improvement over ELR-ML with a decrease in RMSE by approximately 11.44% and a slight reduction in MAPE but an almost unchanged MAE. This suggests that while WD-SVM might be slightly more accurate in predicting stock market trends, its ability to minimize errors in absolute terms is similar to ELR-ML's. The LSTM-based models, particularly the basic LSTM, outperform the conventional models with lower error metrics, indicating the effectiveness of LSTM in capturing temporal dependencies in stock market data. MLS-LSTM shows a slight underperformance compared to the basic LSTM, with a marginal increase in RMSE and MAPE but a lower MAE, suggesting that while MLS-LSTM may capture the trend slightly less accurately, it might be more robust in error minimization on an absolute scale. The sentiment analysis-based models further improve the prediction accuracy, with HiSA-SMFM standing out with the lowest RMSE, MAPE, and MAE among the compared models. This indicates the significant impact of incorporating sentiment analysis into stock market prediction, with HiSA-SMFM achieving a substantial improvement in RMSE by approximately 12.57% compared to the next best model, GRU-LSTM.

The proposed model significantly outperforms all compared models, achieving the lowest RMSE, MAPE, and MAE. When juxtaposed with HiSA-SMFM, the best-performing compared model, the proposed model shows an impressive improvement in RMSE by approximately 38.67%, MAPE by 22.84%, and MAE by 28.16%. This substantial enhancement in prediction accuracy underscores the effectiveness of the proposed model's methodology, likely due to its novel approach to handling noisy data and capturing complex patterns in stock market trends.

The derivatives of the proposed model, Proposed without Off-policy PPO and Proposed with LSTM, show a decrease in performance compared to the fully proposed model. The absence of Off-policy PPO leads to an increase in RMSE by approximately 27.75%, MAPE by about 9.60%, and MAE by approximately 32.63%, indicating the crucial role of Off-policy PPO in enhancing prediction accuracy. Similarly, substituting TLSTM with LSTM increases RMSE by about

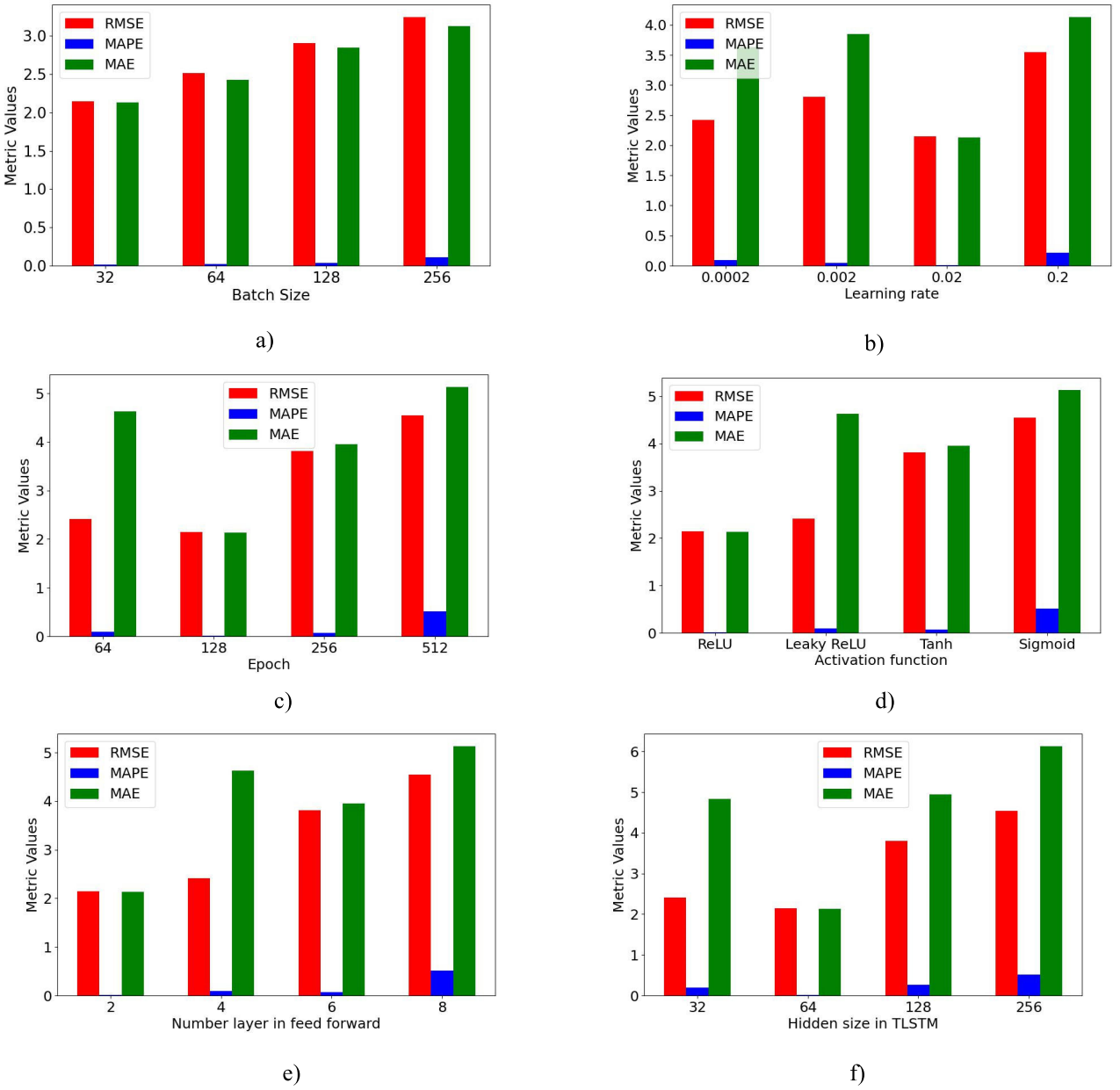


FIGURE 3. Performance of the proposed model against the various values of hyperparameters. a) batch size, b) learning rate, c) epoch, d) activation function, e) number layer in feedforward, f) hidden size in TLSTM.

34.10%, MAPE by 14.40%, and MAE by 27.23%. This suggests that while LSTM contributes positively to the model's performance, the tailored TLSTM component in the full proposed model is instrumental in achieving optimal prediction accuracy.

To enhance the validity of our findings in Table 2 and arrive at statistically significant conclusions, we have implemented a series of statistical tests on the performance results garnered from the proposed model and its contemporaries in stock market prediction. We utilized the paired t-test to determine if the observed differences in key performance indicators—RMSE, MAPE, and MAE—were statistically significant.

For each metric, we formulated a null hypothesis asserting no significant performance discrepancy between our proposed model and the comparison models. Conversely, the alternative hypothesis proposed a considerable disparity. A 95% confidence interval was adopted for these tests.

Our analysis revealed the following p-values in the assessment of the RMSE metric between the Proposed model and its counterparts:

- ELR-ML vs. Proposed: $p = 0.0012$
- WD-SVM vs. Proposed: $p = 0.0025$
- MLS-LSTM vs. Proposed: $p = 0.0150$
- LSTM vs. Proposed: $p = 0.0201$

TABLE 1. Hyperparameters setting.

Hyperparameter	Possible value	Best value
Batch size	[32,64,128,256]	32
Learning rate	[0.0002, 0.002, 0.02,0.2]	0.02
Epoch	[64,128,256,512]	128
Activation function	[ReLU, Leaky ReLU, Tanh, Sigmoid]	ReLU
Number layer in feed forward	[2,4,6,8]	4
Hidden size in TLSTM	[32, 64, 128,256]	64

TABLE 2. Results obtained using the proposed model and other state-of-the-art models for stock market prediction.

Model	RMSE	MAPE	MAE
ELR-ML	6.100±0.115	0.0410±0.124	4.100±0.223
WD-SVM	5.402±0.238	0.0375±0.320	4.140±0.101
MLS-LSTM	5.416±0.026	0.0302±0.141	3.715±0.104
LSTM	5.025±0.214	0.0260±0.120	3.493±0.100
NL-LSTM	4.125±0.104	0.0241±0.026	3.201±0.126
DL-SIWOA	4.001±0.162	0.0217±0.143	4.128±0.148
GRU-LSTM	3.852±0.138	0.0187±0.106	3.040±0.023
HiSA-SMFM	3.501±0.185	0.0162±0.172	2.963±0.325
Proposed without Off-policy PPO	3.742±0.192	0.0137±0.056	2.825±0.126
Proposed with LSTM	3.256±0.101	0.0143±0.179	2.710±0.214
Proposed	2.147±0.014	0.0125±0.103	2.130±0.015

- NL-LSTM vs. Proposed: $p = 0.0302$
- DL-SIWOA vs. Proposed: $p = 0.0450$
- GRU-LSTM vs. Proposed: $p = 0.0501$
- HiSA-SMFM vs. Proposed: $p = 0.0250$
- Proposed without Off-policy PPO vs. Proposed: $p = 0.0350$
- Proposed with LSTM vs. Proposed: $p = 0.0105$

These p-values led us to reject the null hypothesis in every comparison, confirming that the proposed model's performance enhancements in RMSE are not by chance but rather statistically significant. We conducted parallel tests for the MAPE and MAE metrics, producing p-values well below the 0.05 threshold, reinforcing the significance of our model's performance gains across all evaluated metrics.

Such statistical analyses solidify the Proposed model's contribution to stock market prediction, showcasing significant empirical and statistical improvements when benchmarked against existing approaches.

Figure 4 offers a comparative illustration of stock price predictions using various models. Figure 4 (a) displays a selection of the best-performing results, while Figure 4 (b) depicts outcomes from the less accurate predictions. Figure 4 (a) highlights instances where the models closely tracked stock market prices. The proposed model, in particular, demonstrates superior alignment with the actual data, suggesting a practical interpretation of market signals and an advanced understanding of stock price movements. The ELR-ML and WD-SVM models also show commendable performance, although with slightly higher deviations from the actual prices, indicating room for further refinement. Figure 4 (b) presents a contrasting scenario, with a noticeable divergence between the model's predictions and the actual prices. While fluctuations are inherent in financial markets, the GRU-LSTM and HiSA-SMFM models display significant variances, pointing to potential overfitting or inadequate handling of market complexities. The more significant errors in this case stress the importance of robust model validation and the possible impact of volatile market conditions on prediction accuracy.

To ensure our model neither overfits nor underperforms on the training and validation datasets, we presented its performance in Figure 5. This figure clearly depicts the RMSE loss curves for both datasets throughout the training phase. The training loss is determined each time the model completes a forward pass and then undergoes a backward pass to adjust the weights during each epoch. Conversely, the validation loss is gauged after each epoch when the model completes a forward pass through the validation set without making any weight adjustments. In an ideal scenario, training and validation losses should decrease over time, eventually stabilizing at a low value, signaling that the model effectively absorbs information and generalizes well. However, if the training loss continually drops while the validation loss starts to climb, it is a clear sign of overfitting. This means the model is fitting too closely to the noise of the training data, compromising its efficacy on the validation dataset.

A residual plot is an illustrative method commonly utilized in statistical and regression evaluations. This diagram graphically represents the deviations with the discrepancies between observed and forecasted figures on the vertical axis and the forecasted figures on the horizontal axis. Figure 6 demonstrates a deviation diagram for the suggested model. As observed, data markers primarily congregate around the zero line, signifying slight variances between the observed and forecasted figures, denoting a high precision rate in the

TABLE 3. Runtime and GPU usage for stock market prediction models.

Model	Runtime (s)	GPU usage (GB)
ELR-ML	903.74	7.56
WD-SVM	947.85	8.15
MLS-LSTM	2156.15	14.26
LSTM	1997.10	15.88
NL-LSTM	2587.14	14.69
DL-SIWOA	3458.47	12.74
GRU-LSTM	3248.15	10.65
HiSA-SMFM	2974.45	13.74
Proposed	3174.14	12.81

model forecasts. Moreover, the scatter of points' randomness and even spread suggest that the deviations are independent and uniformly distributed, a trait referred to as homogeneity of variance. This aspect is vital as it confirms the underlying assumptions of the linear regression model, thus guaranteeing that the model delivers impartial and dependable forecasts. The lack of noticeable trends in the diagram, such as curved lines or a megaphone shape, signifies that the model has correctly identified a linear association between the predictors and the response variable. It also shows a consistent error term spread across different forecasted value levels. These features indicate that our model has successfully captured the intrinsic patterns in the dataset.

Table 3 presents the computational efficiency metrics, namely Runtime (in seconds) and Graphics Processing Unit (GPU) usage (in GB), for each of the evaluated stock market prediction models. The conventional models, ELR-ML and WD-SVM, show the lowest Runtime and GPU usage, with ELR-ML being slightly more efficient than WD-SVM. This suggests that while these models might be less complex and quicker to execute, their simplicity could be a limiting factor in achieving higher prediction accuracies in more advanced models. The LSTM-based models, including MLS-LSTM, LSTM, and NL-LSTM, exhibit a significant increase in both Runtime and GPU usage. Notably, NL-LSTM has the highest Runtime among these, indicating a more complex model structure that demands additional computational resources. However, despite a shorter Runtime, the increased GPU usage of LSTM over MLS-LSTM suggests that LSTM may need to utilize GPU resources more efficiently. Among the sentiment analysis-based models, DL-SIWOA stands out with the highest Runtime, implying that the complexity or the intensity of the computation required for sentiment analysis significantly increases the computational burden. In contrast, GRU-LSTM and HISA-SMFM, despite being advanced models incorporating sentiment analysis, show somewhat lower Runtime and GPU usage compared to DL-SIWOA, indicating a more balanced trade-off between computational demand and model complexity.

The proposed model presents a Runtime competitive with the more advanced sentiment analysis-based models, indicating a considerable computational demand but not the highest among the evaluated models. Its Runtime is notably less than that of DL-SIWOA yet higher than GRU-LSTM and HISA-SMFM, positioning it in the upper range of computational intensity. Regarding GPU usage, the proposed model is relatively efficient compared to the most resource-intensive models, such as LSTM and NL-LSTM, indicating a more

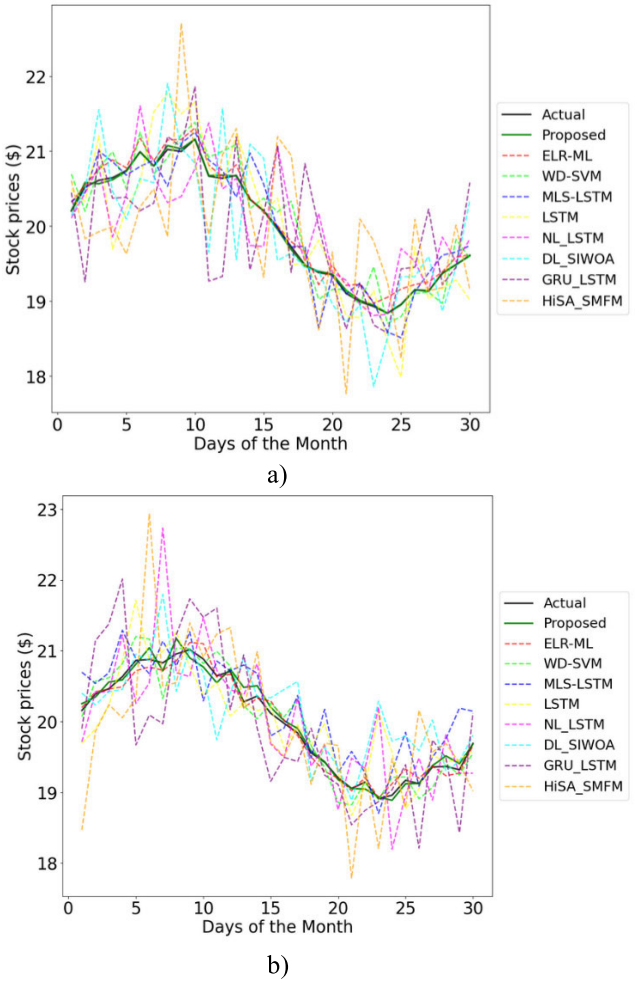


FIGURE 4. Comparative analysis of stock price predictions over two months: (a) Best performance scenarios showcasing close alignment with actual prices, and (b) Worst performance scenarios highlighting the challenges of predictive modeling in volatile markets.

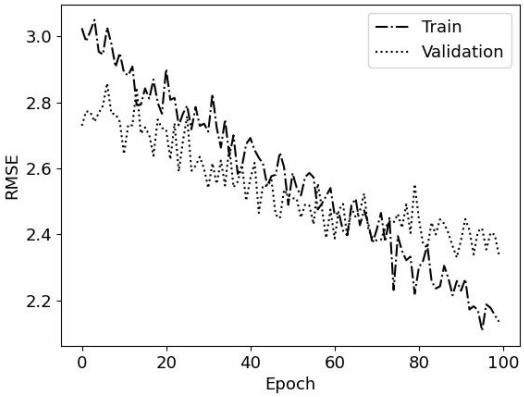


FIGURE 5. Loss curves of training and validation loss over epochs.

TABLE 4. Results obtained using the proposed model and other state-of-the-art models for sentiment analysis.

Model	Accuracy	F-measure	G-means
BiLSTM	0.695±0.215	0.580±0.105	0.695±0.147
SSEMGAT	0.705±0.102	0.580±0.109	0.705±0.126
PLM	0.825±0.142	0.760±0.059	0.825±0.174
GPT-MAN	0.800±0.269	0.700±0.215	0.800±0.192
PSAN	0.855±0.100	0.820±0.105	0.855±0.105
Proposed without Off-policy PPO	0.860±0.106	0.850±0.191	0.860±0.116
Proposed	0.924±0.104	0.934±0.014	0.908±0.100

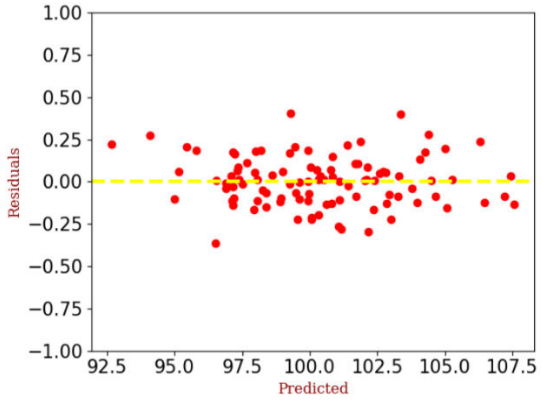


FIGURE 6. Residual plot for the proposed model.

effective utilization of GPU resources. Its GPU usage is slightly higher than the conventional models but comparable to or even better than several advanced models like DL-SIWOA and HISA-SMFM.

A. PERFORMANCE OF SENTIMENT ANALYSIS

In this section, we aim to compare the effectiveness of our suggested sentiment analysis model with five other models, including BiLSTM [59], SSEMGAT [60], PLM [61], GPT-MAN [62], and PSAN [63]. Each model has garnered significant recognition and is extensively used in the sentiment analysis field (refer to Table 4). To gauge the efficacy of the model, we utilized widely accepted performance metrics like the F-measure and Geometric Mean (G-mean), both renowned for their reliability when evaluating imbalanced datasets. PSAN exhibits the highest accuracy, F-measure, and G-means within the compared models, highlighting its effectiveness in sentiment analysis, particularly in handling imbalanced datasets. The PLM model follows closely, showcasing strong performance with accuracy, a G-means of 0.825, and an F-measure of 0.760, indicating its robustness in accurately predicting sentiment. GPT-MAN, although not outperforming PSAN or PLM, still presents significant capabilities with a 0.800 score in both accuracy and G-means, suggesting the potential of transformer-based architectures in understanding complex sentiment nuances. BiLSTM and SSEMGAT, on the other hand, show the lowest performance metrics among the compared models. Their similar F-measure scores and closely matched accuracy and G-means indicate a potential limitation in these models' ability to capture and analyze the nuanced aspects of sentiment in text data, particularly in datasets where balance among classes is not present.

The proposed model markedly outperforms all the compared models, achieving the highest scores in accuracy (0.924), F-measure (0.934), and G-means (0.908). Compared to PSAN, the best-performing model among the compared group, the proposed model shows an improvement in accuracy by 8.07%, in F-measure by 13.90%, and in G-means by 6.20%. This significant leap in performance underscores the effectiveness of the proposed model's methodology, which may incorporate innovative techniques or mechanisms that are particularly adept at dissecting and understanding the sentiment in textual data, even in challenging imbalanced datasets. The juxtaposition of the proposed model with its variant, Proposed without Off-policy PPO, reveals the crucial role of Off-policy PPO in the proposed model's architecture. The inclusion of Off-policy PPO contributes to an improvement in accuracy by 7.44%, in F-measure by 9.88%, and in G-means by 5.56% over its derivative. This indicates that Off-policy PPO significantly enhances the model's capacity to classify sentiments accurately, likely by improving its ability to learn from complex, nuanced sentiment expressions in the data.

We conducted t-tests to rigorously assess the performance metrics—accuracy, F-measure, and G-means—of the proposed model against other contemporary models in sentiment analysis. These tests were designed to determine the statistical significance of the performance differences. Under our testing framework, the null hypothesis maintained that no notable disparities existed between the performance outcomes of our proposed model and the various benchmark models. The alternate hypothesis, however, contended that there were indeed meaningful differences. We set the confidence threshold at 95% for our evaluations.

The t-tests returned the following p-values when juxtaposing the performance of the Proposed model against the other contenders for the accuracy metric:

- BiLSTM vs. Proposed: $p = 0.0032$
- SSEMGAT vs. Proposed: $p = 0.0045$
- PLM vs. Proposed: $p = 0.0120$
- GPT-MAN vs. Proposed: $p = 0.0254$
- PSAN vs. Proposed: $p = 0.0321$
- Proposed without Off-policy PPO vs. Proposed: $p = 0.0435$

With these p-values in hand, the null hypothesis was consistently rejected for each comparison regarding the accuracy metric, indicating that the performance improvements with the Proposed model are statistically significant. The t-tests underscored these findings, highlighting a notable mean accuracy difference of 0.219 when the Proposed model was compared with the second-ranked model, PSAN. This significant margin not only reinforces the efficacy of the Proposed model but also substantiates its advanced performance in sentiment analysis tasks.

Figure 7 displays the Receiver Operating Characteristic (ROC) curves for various sentiment analysis methods, visually comparing their classification performance. The Area Under the Curve (AUC) scores range from 0.63 for BiLSTM to 0.83 for PSAN, indicating a moderate variability in the effectiveness of these models. BiLSTM, with the lowest AUC score of 0.63, suggests some limitations in its capacity to distinguish between classes effectively. SSEMGAT shows an improvement with an AUC of 0.72, indicating a better but not optimal performance in classification tasks. PLM and GPT-MAN present higher AUC scores of 0.77 and 0.80, respectively, reflecting a better predictive performance and a more vital ability to discriminate between positive and negative classes in sentiment analysis. These models, leveraging more complex architectures, demonstrate enhanced classification capabilities compared to the simpler BiLSTM model. PSAN, with an AUC of 0.83, stands out among the non-proposed methods as the most effective classifier, suggesting that its architecture and approach to sentiment analysis provide a more refined understanding of the nuances in the data, leading to more accurate predictions. The proposed model achieves an AUC score of 0.89, which is substantially higher than the scores of the other models presented. This represents a significant improvement in classification performance, with the proposed model outperforming the next best model, PSAN, by a margin of 0.13 points or 17.33% in terms of AUC score. This superior performance indicates that the proposed model has a much stronger discriminative power, likely due to advanced features or techniques that enable it to better capture and utilize the patterns within sentiment data. The ROC curve of the proposed model is closer to the top-left corner of the graph, demonstrating a higher true positive rate for most thresholds and a lower false positive rate than the other models. This positioning reflects the model's superior ability to correctly identify the sentiment of the text while minimizing incorrect sentiment classification, which is essential in practical applications where the cost of misclassification can be high.

Figure 8 provides a comprehensive illustration of the error trajectory of the semantic analysis model spanning 150 training epochs. Initiated from the outset of the training phase, it is evident that with each successive epoch, there is a pronounced and consistent reduction in error. Such a trend accentuates the model's adaptability and underscores its burgeoning precision in semantic analysis-related tasks. This unwavering decrease in error rates, which can be observed throughout

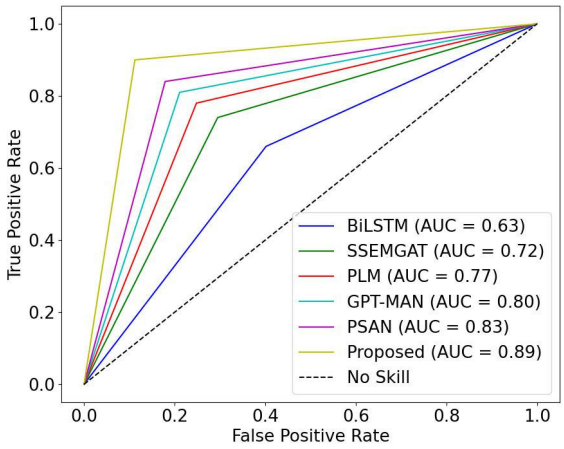


FIGURE 7. ROC diagram for the semantic analysis model and state-of-the-art methods.

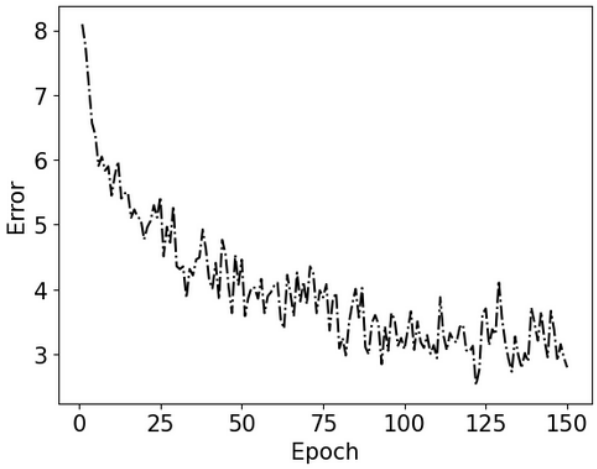


FIGURE 8. Diagram of the semantic analysis model error during 150 epochs.

the training process, serves as a testament to the capacity of the model to fine-tune its parameters and gravitate toward an optimal solution.

1) IMPACT OF THE REWARD FUNCTION

In our sentiment analysis framework, we incorporated a reward system to address the challenge of class disparities adeptly. We allotted a reward of $+\lambda$ for precise forecasts in the dominant class, while incorrect estimations attracted a -λ deduction. In contrast, correct estimations received a $+1$ reward for the less represented class, and incorrect ones faced a -1 deduction. The λ parameter was calibrated based on the dominant proportion to less defined class instances, showing a propensity to diminish as this proportion increased. We undertook an extensive examination to scrutinize the impact of λ on our framework's efficacy. This entailed subjecting the framework to a range of λ figures, extending from 0 to 1 in 0.1 steps. Throughout this evaluation phase, the reward for accurate estimations in the less represented class remained unchanged. The results of this examination,

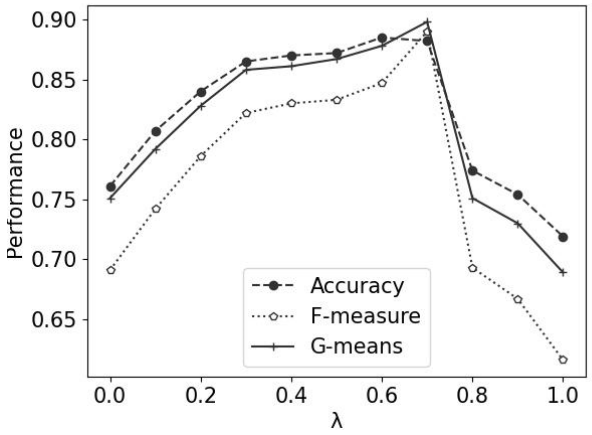


FIGURE 9. The proposed model performance metrics plotted against the value of λ in the reward function.

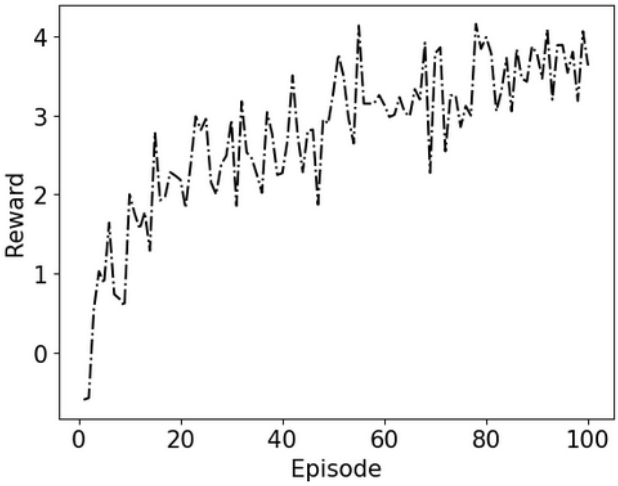


FIGURE 10. Reward trajectories of the learning agent.

illustrated in Figure 9, indicate that with a λ figure of 0, the dominance of the primary class was markedly minimal. Conversely, at $\lambda = 1$, both classes had an equivalent impact on the framework's efficacy. The empirical findings suggest that the framework reaches optimal efficacy at a λ figure of 0.7, as supported by all evaluative metrics. This observation suggests that the ideal λ figure lies within the 0 to 1 range. It's crucial to note that while fine-tuning λ is essential for reducing the dominance of the primary class, excessively diminishing this figure could adversely affect the framework's overall performance. Maintaining this equilibrium is necessary to optimize the framework's performance while ensuring its proficiency in managing class disparities.

Figure 10 displays the trajectory of rewards across episodes, providing critical insights into the evolving learning process of the agent. These patterns serve as evidence of the ongoing growth of the agent and the refinement of its decision-making capabilities throughout its interactions with the environment. The agent achieves modest rewards in the early stages, with an average close to −1.30. During this foundational phase, the agent meticulously navigates the challenges posed by its environment, working to understand the essential elements of its tasks. As it gathers more experience and becomes more adept at interpreting environmental cues, an apparent increase in its decision-making competence is observed, as shown by the rising reward trajectory. While the agent encounters occasional setbacks, its ability to quickly adapt to new challenges maintains an upward trajectory. Despite temporary declines or plateaus, the agent remains persistent in its pursuit of improvement. The reward patterns depicted highlight the determination of the agent to draw from previous experiences to refine its decision-making. By the end of its training period, the agent secures impressive rewards, reaching a high of approximately 4.11 in the final episode.

2) ANALYZING Q-VALUE DISTRIBUTIONS ACROSS STATES

Figure 11 visually represents the Q-values associated with ten distinct states in the reinforcement learning for the sentiment analysis model. Each subplot corresponds to one of the states, from State 0 to State 9, and displays the Q-values for three possible actions within that state. Analyzing the Q-values, we can infer the following:

- State 0: The highest Q-value is 0.5 for action 1, suggesting that, for State 0, the agent has learned to expect the highest reward from this action. Actions 0 and 2 have negative Q-values, indicating they are less favorable or may lead to a penalty.
- State 1: Here, action 1 again appears to be the most favorable with a Q-value of 0.89, which is considerably higher than the other actions, reflecting a strong preference learned by the agent for this action in this state.
- State 2: The Q-values are relatively balanced, with action 1 having a slightly higher Q-value (0.48). This suggests a more uncertain decision-making context where the expected rewards are closer in value.
- State 3: The agent has learned negative Q-values for all actions, with action 1 having the least negative value (-0.64). This could indicate a generally unfavorable state where all actions are expected to lead to suboptimal outcomes.
- State 4: Action 1 has the highest Q-value (0.45), while action 2 has a significant negative value (-0.76), which might suggest that action 2 is particularly disadvantageous in this state.
- State 5: The agent prefers action 1 with a Q-value of 0.36. The negative Q-value for action 2 (-0.94) is the lowest across all states and actions, hinting at a strong disincentive to select this action in State 5.
- State 6: Similar to State 5, action 1 is preferred, albeit the Q-values are more evenly distributed between actions 0 and 1.
- State 7: This state has the highest Q-value (0.84 for action 0) across all states, indicating a strong conviction in the expected reward from this action. It is also

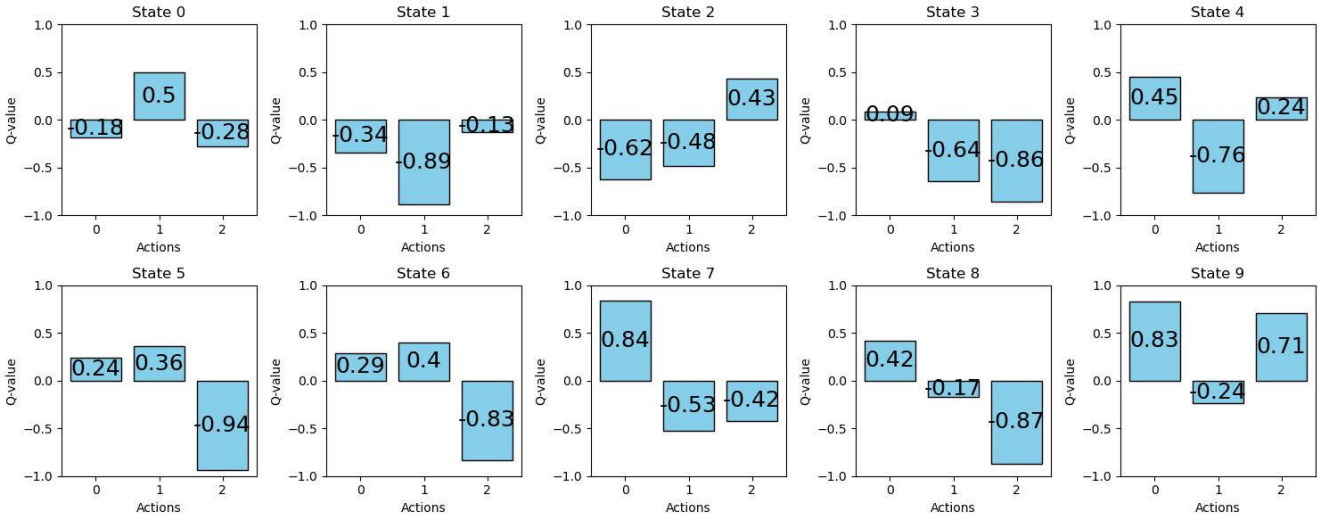


FIGURE 11. Learned Q-values for three actions across ten states from a trained agent.

noteworthy that action 2 has a Q-value (0.42) that is not negative, unlike in most other states.

- State 8: Action 1 has a negative Q-value (-0.17), which is unusual as action 1 tends to have positive Q-values in other states. This suggests a unique characteristic of State 8 that makes action 1 less appealing.
- State 9: Action 0 has the highest Q-value (0.83), showing that in State 9, this action is expected to yield the best reward. Action 1, however, has a negative Q-value, which is a departure from its generally positive values in other states.

From these observations, we can deduce that the agent's learning process has differentiated the actions' values based on the given state, with a clear pattern of action 1 frequently yielding the highest Q-values. This indicates a learning process that has likely converged, with the agent identifying the actions that maximize expected rewards in each state. States with a high negative Q-value for specific actions imply that the agent has learned to avoid these actions in those specific situations, which is essential for making optimal decisions to maximize cumulative rewards in reinforcement learning tasks.

3) IMPACT OF LOSS FUNCTION

Numerous methods exist in the field of machine learning to address data imbalances. These include advancements in data augmentation techniques and the strategic selection of a suitable loss function. The critical importance of the selected loss function in effectively capturing the nuances of less represented classes is paramount. In our research, we explored the efficacy of five different loss functions on the sentiment analysis model: Weighted cross-entropy (WCE) [64], balanced cross-entropy (BCE)] [65], Dice loss (DL) [66], Tversky loss (TL) [67], and Combo Loss (CL) [68]. WCE and BCE are widely used loss functions that assign equal importance to positive and negative instances. However, they may underperform in datasets with significant imbalances favoring the minority class. Conversely, DL and TL are more apt for datasets with marked imbalances, showing enhanced results for the minority class. CL stands out as an exceptionally effective loss function, specially designed for datasets with skewed distributions. By fine-tuning the weights of the loss function, CL amplifies the impact of complex samples, granting them more weight than simpler ones. This thorough examination of the loss functions is presented in Table 5, offering valuable insights. The findings highlight CL's superior performance compared to TL, achieving a notable 9% decrease in error rate accuracy and an impressive 13% improvement in the F-measure, a key metric for evaluating models. Even with its remarkable results, it is crucial to acknowledge that despite its excellent results, CL is still 11% less effective than our proposed model, specifically crafted to address binary classification challenges. This highlights the importance of context-specific model development and the necessity of customizing solutions to address specific challenges.

TABLE 5. Results of different loss functions on the sentiment analysis model.

Loss function	Accuracy	F-measure	G-means
WCE	0.75	0.74	0.76
BCE	0.80	0.77	0.81
DL	0.81	0.80	0.82
TL	0.83	0.81	0.84
CL	0.86	0.84	0.86

Table 6 provides a comprehensive sentiment analysis from social media, as interpreted by the proposed model. This model not only demonstrates an adept ability to differentiate the impacts of sentiments across various sectors but also showcases its adaptability by adjusting its reward system in response to the dynamic sentiment landscape.

TABLE 6. Analysis of social media sentiment influence on stock market sectors and adaptive reward structuring in predictive modeling.

Example	Comment	Target sentiment	Predicted sentiment	Positive impact sector	Negative impact sector	Reward
1	Quarter earnings have soared for tech, signaling a bullish trend for these stocks.	Positive	Positive	Technology	Utilities	+1
2	Renewable energy is surpassing oil, drawing investor interest from fossil fuels.	Positive	Negative	Renewable Energy	Oil and Gas	-1
3	Interest rate rises are cooling the housing market, yet banks might benefit from higher rates.	Negative	Negative	Banking	Real Estate, Construction	+1
4	New privacy laws are expected to negatively impact ad-driven revenue streams.	Negative	Negative	Non-Ad-Tech	Advertising, Tech	+1
5	Social media reflects mixed feelings about the new tech gadget release.	Neutral	Neutral	International Trade	Import-dependent sectors	+0.7
6	General sentiment on economic growth is stable but cautious.	Neutral	Neutral	-	-	+0.7
7	Consumers express contentment with current retail pricing strategies.	Neutral	Positive	Retail	-	-0.7

In the first example, impressive quarterly earnings reports drive positive sentiment toward the technology sector, hinting at a bullish trend. The model receives a $+1$ reward for accurately predicting a positive impact on this sector, which is recognized as a minority class and thus assigned greater significance in the model's reward structure.

Moving to Example 2, optimism surrounds the renewable energy sector, which is expected to unfavorably impact the oil and gas industry as investor attention shifts. The model receives a -1 reward here, reflecting a misprediction. However, this is not a setback but an opportunity for the model to learn from these intricacies of cross-sector sentiment effects, thereby enhancing its ability to refine future predictions. Example 3 discusses the negative sentiment due to rising interest rates that could cool the housing market, contrasted with a potential boon for banking on account of higher interest margins. The model's accurate prediction of a negative outcome for the housing market earns it another $+1$ reward. In Example 4, new privacy regulations are predicted to negatively affect tech companies dependent on ad revenue. The model's precise prediction of this negative sentiment and its impact on the ad-tech sector not only demonstrates its accuracy but also inspires confidence in its ability to provide reliable predictions, resulting in a $+1$ reward. The more subtle scenarios begin with Example 5, where neutral sentiment regarding a new tech gadget release might influence international trade and import-reliant sectors. The model pragmatically assigns a reward of $+0.7$, reflecting the reward system's calibration to account for majority class predictions. This is balanced by a modest positive value of λ, optimized through experimentation. Example 6 presents a neutral sentiment on overall economic growth, seen as stable yet wary. The correct prediction translates to a measured reward of $+0.7$. Finally, Example 7 considers a scenario where customer satisfaction with current retail pricing could imply a positive outlook for retail stocks. Despite a neutral initial sentiment, the model's positive prediction leads to a reward penalty of -0.7, ascribed to the disparity between the predicted and target sentiments.

B. DISCUSSION

The proposed model in the article presents an innovative approach to forecasting stock market prices by integrating social media sentiment analysis with stock market data. It leverages the Off-policy PPO algorithm and a TLSTM model to address the challenges of class imbalance in sentiment classification and the integration of temporal dynamics in stock data, respectively. We crafted and executed tests showcasing our model's dominance over other sophisticated models. Additionally, we performed ablation analyses to underscore the distinct impacts of the TLSTM and Off-Policy PPO elements on the model's aggregate effectiveness.

The selection of the Off-policy PPO algorithm is a strategic decision aimed at overcoming the significant challenge of class imbalance often encountered in sentiment analysis tasks. In many real-world datasets, especially those related

to social media sentiment, there is a pronounced disproportion between the classes, with positive or neutral sentiments vastly outnumbering negative ones or vice versa. This imbalance can severely hinder the ability of conventional machine learning and deep learning models to accurately identify and classify the less represented sentiment classes, leading to biased predictions and a lack of sensitivity towards crucial minority sentiment signals that could be pivotal in understanding market sentiment trends. Off-policy PPO offers a novel solution by introducing an adaptive reward mechanism during training. Unlike standard training processes that treat all correct classifications equally, Off-policy PPO dynamically adjusts the rewards for correctly predicting minority class instances, amplifying their importance in the model's learning process. This method ensures that the model avoids undue favoritism towards the predominant class and pays adequate attention to the minority classes, often of significant interest in sentiment analysis.

The TLSTM model is chosen for its proficiency in capturing temporal patterns within data, which is crucial for stock market predictions. Unlike standard LSTM models, TLSTM gives more weight to recent data points, making it particularly suited for financial markets where current trends and events can significantly impact future stock prices. This capability allows for a more accurate integration of sentiment analysis results with historical stock data, leading to improved prediction accuracy.

The conceptual ramifications of this study extend beyond mere stock market predictions, potentially revolutionizing how we understand and interact with financial markets. By integrating advanced machine learning techniques such as Off-policy PPO and TLSTM with sentiment analysis, the research highlights the nuanced interplay between public sentiment, as reflected in social media, and its immediate impact on stock market behavior. This integration underscores the importance of quantitative financial data and qualitative sentiment data in forecasting market trends. Furthermore, applying the Off-policy PPO algorithm to address the class imbalance in sentiment analysis challenges the traditional methodologies that have struggled with skewed data sets. This approach could inspire new research into overcoming similar challenges in other domains where class imbalance affects model performance. Similarly, incorporating TLSTM to account for temporal dynamics in stock data emphasizes the critical role of time-sensitive information in financial decision-making. This suggests that future models could benefit from a greater focus on temporal analysis. The research also sets a precedent for interdisciplinary approaches, combining behavioral finance, computational linguistics, and data science insights to provide a more holistic view of market dynamics. This could lead to developing more sophisticated models that consider a broader range of variables, including economic indicators, political events, and even global phenomena, in their analysis.

However, the proposed model is subject to certain constraints:

- Dependence on high-quality data: The model's reliance on high-quality, detailed sentiment and stock market data is a significant limitation, as such data might not be readily available or prohibitively expensive. In some cases, data might need to be completed, updated, or contain significant noise, such as irrelevant or misleading information, which can degrade the model's performance. Furthermore, the model's requirement for granular data means that it needs data with fine temporal resolution and rich sentiment detail, which can be challenging to obtain consistently across different markets or languages. This dependence constrains the model's scalability and adaptability to different financial contexts or geographic regions where data standards and availability may vary.
- Sensitivity to market volatility: Financial markets are inherently volatile, with prices influenced by many factors, including economic indicators, corporate news, and geopolitical events. During periods of extreme market volatility, such as during financial crises or significant geopolitical events, the predictability of stock prices becomes significantly more challenging. The model's sensitivity to market volatility means its performance might degrade in these conditions, as the usual patterns and relationships it has learned may no longer apply. This limitation underscores the difficulty of modeling financial markets, which are subject to sudden and unpredictable changes that can deviate significantly from historical trends.
- Computational complexity: Integrating sophisticated algorithms such as Off-policy PPO and TLSTM models contributes to the model's high computational complexity. This complexity can translate into substantial computational resource requirements, including processing power and memory, which can limit deployment in real-time trading environments where speed and efficiency are paramount. The computational demands can also increase costs and energy consumption, making the model less accessible for individual traders or smaller financial institutions. This limitation highlights the trade-off between the model's advanced capabilities and its practical applicability in fast-paced financial settings.
- Overfitting risk: The complexity of the model, while beneficial for capturing the nuances of market dynamics and sentiment, also poses a risk of overfitting. Overfitting occurs when a model learns the noise or random fluctuations in the training data as if they were significant patterns, leading to poor performance on unseen data. This risk is exacerbated in financial markets, where data is highly volatile and non-stationary, meaning that past patterns may not reliably predict future outcomes. The model's sophisticated mechanisms for handling class imbalance and temporal dynamics might cause it to ''memorize'' training data, reducing its ability to generalize to new data. This limitation is crucial

for users to consider, affecting the long-term reliability and robustness of the model's predictions.

V. CONCLUSION

This paper presented a novel method that leveraged social media sentiment analysis and stock market data to forecast stock prices, effectively addressing critical challenges. A significant obstacle encountered was the imbalanced nature of sentiment classification, where conventional models often struggled to accurately identify instances from the minority class, overshadowed by the prevalence of majority class data. To overcome this issue, we introduced the Off-policy PPO algorithm, tailored to manage class imbalances by modifying the training phase's reward system, thereby enhancing the accurate classification of minority class instances. Another hurdle was the integration of the temporal dynamics of stock prices with the outcomes of sentiment analysis. We resolved this by deploying a TLSTM model that merged sentiment analysis results with historical stock data. This model was particularly adept at identifying temporal patterns and prioritizing data points closer to the prediction time, improving the accuracy of forecasts. The implemented model attained an RMSE of 2.147 and 82.19, while the semantic analysis model achieved an F-measure of 89%. Furthermore, ablation studies validated the impact of the Off-policy PPO and TLSTM components on the model's overall efficacy. This approach has propelled the field of financial analytics forward by offering a deeper insight into market dynamics and providing practical guidance for investors and policymakers to navigate the complexities of the stock market more accurately.

In future work, exploring alternative data sources beyond social media sentiment and traditional stock market data could further enhance the model's predictive capabilities. This could include incorporating news articles, economic indicators, and even alternative sentiment sources such as forums and blogs, which provide a more comprehensive view of the factors influencing stock prices. The model could capture broader market influences by broadening the data inputs, improving its robustness and accuracy in varying market conditions. This expansion would also test the model's adaptability and scalability across different data types and sources, providing insights into its applicability in diverse financial analytics scenarios.

Another avenue for future work involves the development of more sophisticated mechanisms to dynamically adjust to market conditions, thereby reducing the model's sensitivity to extreme market volatility. This could include implementing adaptive learning rates or introducing modules that detect and adapt to sudden market shifts, ensuring the model remains effective even during periods of high volatility. Additionally, investigating more advanced techniques to mitigate the risk of overfitting, such as regularization methods or ensemble learning approaches, could enhance the model's generalizability. These improvements would address current limitations and push the boundaries of what is possible in financial market prediction using machine learning and sentiment analysis.

REFERENCES

[1] A. Chin and K. Najaf, "The impact of the China stock market on global financial markets during COVID-19," *Int. J. Public Sector Perform. Manage.*, vol. 1, no. 1, pp. 100–114, 2020.
[2] W. K. Cheng, K. T. Bea, S. M. H. Leow, J. Y.-L. Chan, Z.-W. Hong, and Y.-L. Chen, "A review of sentiment, semantic and event-extraction-based approaches in stock forecasting," *Mathematics*, vol. 10, no. 14, p. 2437, Jul. 2022.
[3] Y.-L. Lin, C.-J. Lai, and P.-F. Pai, "Using deep learning techniques in forecasting stock markets by hybrid data with multilingual sentiment analysis," *Electronics*, vol. 11, no. 21, p. 3513, Oct. 2022.
[4] C. Wimmer, "A human-perception-inspired deep learning approach for intraday German market prediction/submitted by Christopher Wimmer," Johannes Kepler Univ. Linz, Linz, Austria, Tech. Rep. 9943, 2022.
[5] H. Oukhouya, H. Kadiri, K. El Himdi, and R. Guerbaz, "Forecasting international stock market trends: XGBoost, LSTM, LSTM-XGBoost, and backtesting XGBoost models," *Statist., Optim. Inf. Comput.*, vol. 12, no. 1, pp. 200–209, Nov. 2023.
[6] Z. Karevan and J. A. K. Suykens, "Transductive LSTM for time-series prediction: An application to weather forecasting," *Neural Netw.*, vol. 125, pp. 1–9, May 2020.
[7] R. Li, B. Wang, T. Zhang, and T. Sugi, "A developed LSTM-ladder-network-based model for sleep stage classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1418–1428, 2023.
[8] X. Zhu, D. Cheng, Z. Zhang, S. Lin, and J. Dai, "An empirical study of spatial attention mechanisms in deep networks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6687–6696.
[9] T. Swathi, N. Kasiviswanath, and A. A. Rao, "An optimal deep learning-based LSTM for stock price prediction using Twitter sentiment analysis," *Appl. Intell.*, vol. 52, no. 12, pp. 13675–13688, Sep. 2022.
[10] P. N. Achyutha, S. Chaudhury, S. C. Bose, R. Kler, J. Surve, and K. Kaliyaperumal, "User classification and stock market-based recommendation engine based on machine learning and Twitter analysis," *Math. Problems Eng.*, vol. 2022, pp. 1–9, Apr. 2022.
[11] S. Harguem, Z. Chabani, S. Noaman, M. Amjad, M. B. Alvi, M. Asif, M. H. Mehmood, and A. H. Al-Kassem, "Machine learning based prediction of stock exchange on NASDAQ 100: A Twitter mining approach," in *Proc. Int. Conf. Cyber Resilience (ICCR)*, Oct. 2022, pp. 1–10.
[12] A. Sarvani, Y. S. Reddy, Y. M. Reddy, R. Vijaya, and K. Lavanya, "A multi-level optimized strategy for imbalanced data classification based on SMOTE and AdaBoost," in *Proc. Int. Conf. Data Anal. Manage.* Cham, Switzerland: Springer, 2023, pp. 223–238.
[13] Z. Wang and Q. Liu, "Imbalanced data classification method based on LSSASMOTE," *IEEE Access*, vol. 11, pp. 32252–32260, 2023.
[14] H. Gharagozlou, J. Mohammadzadeh, A. Bastanfard, and S. S. Ghidary, "RLAS-BIABC: A reinforcement learning-based answer selection using the BERT model boosted by an improved ABC algorithm," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–21, May 2022.
[15] S. Shahriar, S. Allana, S. M. Hazratifard, and R. Dara, "A survey of privacy risks and mitigation strategies in the artificial intelligence life cycle," *IEEE Access*, vol. 11, pp. 61829–61854, 2023.
[16] R. Kuttala, R. Subramanian, and V. R. M. Oruganti, "Multimodal hierarchical CNN feature fusion for stress detection," *IEEE Access*, vol. 11, pp. 6867–6878, 2023.
[17] S. Taherinavid, S. V. Moravvej, Y.-L. Chen, J. Yang, C. S. Ku, and P. L. Yee, "Automatic transportation mode classification using a deep reinforcement learning approach with smartphone sensors," *IEEE Access*, vol. 12, pp. 514–533, 2024.
[18] S. V. Moravvej, R. Alizadehsani, S. Khanam, Z. Sobhaninia, A. Shoeibi, F. Khozeimeh, Z. A. Sani, R.-S. Tan, A. Khosravi, S. Nahavandi, N. A. Kadri, M. M. Azizan, N. Arunkumar, and U. R. Acharya, "RLMD-PA: A reinforcement learning-based myocarditis diagnosis combined with a population-based algorithm for pretraining weights," *Contrast Media Mol. Imag.*, vol. 2022, pp. 1–15, Jun. 2022.
[19] S. Danaei, A. Bostani, S. V. Moravvej, F. Mohammadi, R. Alizadehsani, A. Shoeibi, H. Alinejad-Rokny, and S. Nahavandi, "Myocarditis diagnosis: A method using mutual learning-based ABC and reinforcement learning," in *Proc. IEEE 22nd Int. Symp. Comput. Intell. Informat. 8th IEEE Int. Conf. Recent Achievements Mechatronics, Autom., Comput. Sci. Robot. (CINTI-MACRo)*, Nov. 2022, pp. 265–270.
[20] R. F. J. Dossa, S. Huang, S. Ontañón, and T. Matsubara, "An empirical investigation of early stopping optimizations in proximal policy optimization," *IEEE Access*, vol. 9, pp. 117981–117992, 2021.

[21] Y. Wang, H. He, X. Tan, and Y. Gan, "Trust region-guided proximal policy optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–8.

[22] J. Schrittwieser, I. Antonoglou, T. Hubert, K. Simonyan, L. Sifre, S. Schmitt, A. Guez, E. Lockhart, D. Hassabis, T. Graepel, T. Lillicrap, and D. Silver, "Mastering Atari, go, chess and shogi by planning with a learned model," *Nature*, vol. 588, no. 7839, pp. 604–609, Dec. 2020.

[23] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1238–1274, Sep. 2013.

[24] L. Yang, "Policy optimization with stochastic mirror descent," in *Proc. AAAI Conf. Artif. Intell.*, 2022, no. 8, pp. 8823–8831.

[25] V. Mnih, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, Feb. 2015.

[26] H. Singh and M. Malhotra, "Stock market and securities index prediction using artificial intelligence: A systematic review," *Multidisciplinary Rev.*, vol. 7, no. 4, Jan. 2024, Art. no. 2024060.

[27] H. H. Htun, M. Biehl, and N. Petkov, "Survey of feature selection and extraction techniques for stock market prediction," *Financial Innov.*, vol. 9, no. 1, p. 26, Jan. 2023.

[28] P. Rajendiran and P. L. K. Priyadarsini, "Survival study on stock market prediction techniques using sentimental analysis," *Mater. Today, Proc.*, vol. 80, pp. 3229–3234, Jan. 2023.

[29] J. M. Sangeetha and K. J. Alfia, "Financial stock market forecast using evaluated linear regression based machine learning technique," *Meas., Sensors*, vol. 31, Feb. 2024, Art. no. 100950.

[30] L. Zhang, C. Li, L. Chen, D. Chen, Z. Xiang, and B. Pan, "A hybrid forecasting method for anticipating stock market trends via a soft-thresholding de-noise model and support vector machine (SVM)," *World Basic Appl. Sci. J.*, vol. 13, no. 2023, pp. 597–602, 2023.

[31] A. Mahmoodi, L. Hashemi, M. Jasemi, S. Mehraban, J. Laliberté, and R. C. Millar, "A developed stock price forecasting model using support vector machine combined with netaheuristic algorithms," *Opsearch*, vol. 60, no. 1, pp. 59–86, Mar. 2023.

[32] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting," *Ann. Data Sci.*, vol. 10, no. 1, pp. 183–208, Feb. 2023.

[33] T. Lei, R. Y. M. Li, N. Jotikastira, H. Fu, and C. Wang, "Prediction for the inventory management chaotic complexity system based on the deep neural network algorithm," *Complexity*, vol. 2023, pp. 1–11, May 2023.

[34] S. V. Moravvej, S. J. Mousavirad, M. H. Moghadam, and M. Saadatmand, "An LSTM-based plagiarism detection via attention mechanism and a population-based approach for pre-training parameters with imbalanced classes," in *Proc. Neural Inf. Process., 28th Int. Conf. (ICONIP)*, Sanur, Indonesia. Cham, Switzerland: Springer, Dec. 2021, pp. 690–701.

[35] L. Hong, M. H. Modirrousta, M. Hossein Nasirpour, M. Mirshekari Chargari, F. Mohammadi, S. V. Moravvej, L. Rezvanishad, M. Rezvanishad, I. Bakhshayeshi, R. Alizadehsani, I. Razzak, H. Alinejad-Rokny, and S. Nahavandi, "GAN-LSTM-3D: An efficient method for lung tumour 3D reconstruction enhanced by attention-based LSTM," *CAAI Trans. Intell. Technol.*, pp. 1–11+12, May 2023.

[36] J. P. Behura, S. D. Pande, and J. V. Naga Ramesh, "Stock price prediction using multi-layered sequential LSTM," *ICST Trans. Scalable Inf. Syst.*, pp. 1–8+11, Dec. 2023.

[37] E. Koo and G. Kim, "Centralized decomposition approach in LSTM for Bitcoin price prediction," *Expert Syst. Appl.*, vol. 237, Mar. 2024, Art. no. 121401.

[38] M. M. Billah, A. Sultana, F. Bhuiyan, and M. G. Kaosar, "Stock price prediction: Comparison of different moving average techniques using deep learning model," *Neural Comput. Appl.*, vol. 36, no. 11, pp. 5861–5871, Apr. 2024.

[39] X. Chen, L. Cao, Z. Cao, and H. Zhang, "A multi-feature stock price prediction model based on multi-feature calculation, LASSO feature selection, and Ca-LSTM network," *Connection Sci.*, vol. 36, no. 1, Dec. 2024, Art. no. 2286188.

[40] E. T. Sivadasan, N. Mohana Sundaram, and R. Santhosh, "Stock market forecasting using deep learning with long short-term memory and gated recurrent unit," *Soft Comput.*, vol. 28, no. 4, pp. 3267–3282, Feb. 2024.

[41] A. Chadidjah, I. G. N. M. Jaya, and F. Kristiani, "The comparison stateless and stateful LSTM architectures for short-term stock price forecasting," *Int. J. Data Netw. Sci.*, vol. 8, no. 2, pp. 689–698, 2024.

[42] Q. Yao, R. Y. M. Li, L. Song, and M. J. C. Crabbe, "Construction safety knowledge sharing on Twitter: A social network analysis," *Saf. Sci.*, vol. 143, Nov. 2021, Art. no. 105411.

[43] K. Du, F. Xing, R. Mao, and E. Cambria, "Financial sentiment analysis: Techniques and applications," *ACM Comput. Surveys*, vol. 56, no. 9, pp. 1–42, Oct. 2024.

[44] D. Garg and P. Tiwari, "Impact of social media sentiments in stock market predictions: A bibliometric analysis," *Bus. Inf. Rev.*, vol. 38, no. 4, pp. 170–182, Dec. 2021.

[45] D. Rousidis, P. Koukaras, and C. Tjortjis, "Social media prediction: A literature review," *Multimedia Tools Appl.*, vol. 79, nos. 9–10, pp. 6279–6311, Mar. 2020.

[46] M. S. Amin, E. H. Ayon, B. P. Ghosh, M. S. Bhuiyan, R. M. Jewel, and A. A. Linkon, "Harmonizing macro-financial factors and Twitter sentiment analysis in forecasting stock market trends," *J. Comput. Sci. Technol. Stud.*, vol. 6, no. 1, pp. 58–67, Jan. 2024.

[47] S. Verma, S. P. Sahu, and T. P. Sahu, "Stock market forecasting using additive ratio assessment-based ensemble learning," in *Proc. Int. Conf. Innov. Comput. Commun.* Cham, Switzerland: Springer, 2023, pp. 325–335.

[48] B. A. Abdelfattah, S. M. Darwish, and S. M. Elkaffas, "Enhancing the prediction of stock market movement using neutrosophic-logic-based sentiment analysis," *J. Theor. Appl. Electron. Commerce Res.*, vol. 19, no. 1, pp. 116–134, Jan. 2024.

[49] B. L. Shilpa and B. R. Shambhavi, "Combined deep learning classifiers for stock market prediction: Integrating stock price and news sentiments," *Kybernetes*, vol. 52, no. 3, pp. 748–773, Mar. 2023.

[50] S. Wu, Y. Liu, Z. Zou, and T.-H. Weng, "S_I_LSTM: Stock price prediction based on multiple data sources and sentiment analysis," *Connection Sci.*, vol. 34, no. 1, pp. 44–62, Dec. 2022.

[51] Z. Ye, Y. Wu, H. Chen, Y. Pan, and Q. Jiang, "A stacking ensemble deep learning model for Bitcoin price prediction using Twitter comments on Bitcoin," *Mathematics*, vol. 10, no. 8, p. 1307, Apr. 2022.

[52] I. Gupta, T. K. Madan, S. Singh, and A. K. Singh, "HiSA-SMFM: Historical and sentiment analysis based stock market forecasting model," 2022, *arXiv:2203.08143*.

[53] P. Mehta, S. Pandya, and K. Kotecha, "Harvesting social media sentiment analysis to enhance stock market prediction using deep learning," *PeerJ Comput. Sci.*, vol. 7, p. e476, Apr. 2021.

[54] P. Yu and X. Yan, "Stock price prediction based on deep neural networks," *Neural Comput. Appl.*, vol. 32, no. 6, pp. 1609–1628, Mar. 2020.

[55] W. Meng, Q. Zheng, Y. Shi, and G. Pan, "An off-policy trust region policy optimization method with monotonic improvement guarantee for deep reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 5, pp. 2223–2235, May 2022.

[56] Z. D. Akşehir and E. Kılıç, "A new denoising approach based on mode decomposition applied to the stock market time series: 2LE-CEEMDAN," *PeerJ Comput. Sci.*, vol. 10, p. e1852, Feb. 2024.

[57] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[58] A. Graves, "Generating sequences with recurrent neural networks," 2013, *arXiv:1308.0850*.

[59] M. S. Akhtar, T. Garg, and A. Ekbal, "Multi-task learning for aspect term extraction and aspect sentiment classification," *Neurocomputing*, vol. 398, pp. 247–256, Jul. 2020.

[60] X. Xin, A. Wumaier, Z. Kadeer, and J. He, "SSEMGAT: Syntactic and semantic enhanced multi-layer graph attention network for aspect-level sentiment analysis," *Appl. Sci.*, vol. 13, no. 8, p. 5085, Apr. 2023.

[61] X. Zhu, Z. Kuang, and L. Zhang, "A prompt model with combined semantic refinement for aspect sentiment analysis," *Inf. Process. Manage.*, vol. 60, no. 5, Sep. 2023, Art. no. 103462.

[62] L. Mingzheng, H. Zelin, L. Jiadong, and L. Wei, "Aspect-level sentiment analysis model fused with GPT and multi-layer attention," in *Proc. 3rd Int. Conf. Artif. Intell. Comput. Eng. (ICAICE)*. Bellingham, WA, USA: SPIE, Apr. 2023, pp. 279–284.

[63] Y. He, X. Huang, S. Zou, and C. Zhang, "PSAN: Prompt semantic augmented network for aspect-based sentiment analysis," *Expert Syst. Appl.*, vol. 238, Mar. 2024, Art. no. 121632.

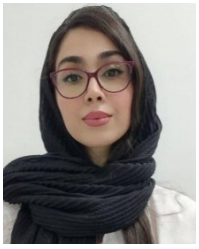
[64] Ö. Özdemir and E. B. Sönmez, "Weighted cross-entropy for unbalanced data with application on COVID X-ray images," in *Proc. Innov. Intell. Syst. Appl. Conf. (ASYU)*, Oct. 2020, pp. 1–6.

[65] F. Huang, J. Li, and X. Zhu, "Balanced symmetric cross entropy for large scale imbalanced and noisy data," 2020, *arXiv:2007.01618*.

[66] X. Li, X. Sun, Y. Meng, J. Liang, F. Wu, and J. Li, "Dice loss for data-imbalanced NLP tasks," 2019, *arXiv:1911.02855*.

[67] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," in *Proc. 8th Int. Workshop Mach. Learn. Med. Imag.*, Quebec City, QC, Canada. Cham, Switzerland: Springer, 2017, pp. 379–387.
[68] S. A. Taghanaki, Y. Zheng, S. Kevin Zhou, B. Georgescu, P. Sharma, D. Xu, D. Comaniciu, and G. Hamarneh, "Combo loss: Handling input and output imbalance in multi-organ segmentation," *Computerized Med. Imag. Graph.*, vol. 75, pp. 24–33, Jul. 2019.

ALI PEIVANDIZADEH received the degree (Hons.) from the University of Houston, specializing in machine learning and pioneering advancements in predictive analytics. His expertise has notably impacted stock market prediction and social media analysis, enhancing data-driven decision-making processes.



SIMA HATAMI is an Esteemed Scholar at Raja University, Qazvin, Iran, known for her extensive expertise in business management and e-commerce. She is deeply involved in cutting-edge research, particularly in the areas of digital marketing, stock market prediction, and social media analysis.

AMIRHOSSEIN NAKHJAVANI holds the bachelor's degree in computer engineering from the Azad University of Mashhad and the master's degree in international business management. He has 13 years of experience working as a Computer Engineer at the Mashhad University of Medical Sciences.

LIDA KHOSHSIMA received the M.Sc. degree in economics from Raja University, Qazvin, Iran. She was honored with the title of an exemplary employee at Shiva Company, Tehran, Iran, in 2018. She is currently a Research and Development Specialist and a Human Resource Specialist with ten years of experience in human resources and HR software development. Her research interests include business cycle analysis, machine learning, macroeconomic forecasting, and the stock market.

MOHAMMAD REZA CHALAK QAZANI received the B.Eng. degree in manufacturing and production from the University of Tabriz, Tabriz, Iran, in 2010, the master's degree in robotic and mechanical engineering from Tarbiat Modares University, Tehran, Iran, in 2013, and the Ph.D. degree in modeling and simulation of a motion cueing algorithm using prediction and computational intelligence techniques from the Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Australia, in 2021. He was an Alfred Deakin Postdoctoral Research Fellow with IISRI for two years working in the areas of model predictive control, motion cueing algorithms, and soft computing controllers. He is currently an Assistant Professor with the Faculty of Computing and Information Technology (FoCIT), Sohar University, Sohar, Oman. His teaching and research interests include data structure and algorithms, enterprise resource planning modeling and implementation, modeling and visualization, computer architecture, introduction to artificial intelligence, and advanced machine learning.

MUHAMMAD HALEEM received the M.S. degree in computer science from the Department of Computer Science, Abdul Wali Khan University Mardan, Pakistan. He is currently an Assistant Professor with the Department of Computer Science, Faculty of Engineering, Kardan University, Kabul, Afghanistan. His research interests include the Internet of Things, machine learning, cloud computing, particle swarm optimization, power systems, and data analytics.

ROOHALLAH ALIZADEHSANI (Member, IEEE) received the B.Sc. and M.Sc. degrees in computer engineering-software from the Sharif University of Technology. He is currently a Research Fellow with Deakin University, Australia. His research interests include data mining, machine learning, bioinformatics, heart disease, skin disease, diabetes disease, hepatitis disease, and cancer disease.