

# A Two-Stage Framework for Stock Price Prediction: LLM-Based Forecasting with Risk-Aware PPO Adjustment

Qizhao Chen

Graduate School of Information Science, University of Hyogo, Kobe, Japan

Email: af24o008@guh.u-hyogo.ac.jp

**How to cite this paper:** Chen, Q.Z. (2025) A Two-Stage Framework for Stock Price Prediction: LLM-Based Forecasting with Risk-Aware PPO Adjustment. *Journal of Computer and Communications*, 13, 120-139.  
<https://doi.org/10.4236/jcc.2025.134008>

**Received:** March 23, 2025

**Accepted:** April 24, 2025

**Published:** April 27, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Accurate prediction of stock prices remains a fundamental challenge in financial markets, with substantial implications for investment strategies and decision making. Although machine learning and deep learning models have significantly advanced the prediction of stock price movements, they often overlook the critical aspect of financial risk. This research proposes a novel framework that integrates Large Language Models (LLMs) with Proximal Policy Optimization (PPO), a reinforcement learning technique, to improve stock price predictions while incorporating risk-adjusted mechanisms. The LLM provides initial predictions based on historical stock data and financial news sentiment, while PPO refines these predictions by adjusting them according to financial risk metrics such as Value at Risk (VaR) and Conditional Value at Risk (CVaR). The framework aims to improve prediction accuracy and provide more reliable forecasts that account for market volatility. Experimental results demonstrate that the proposed LLM-PPO framework outperforms traditional prediction models in terms of both prediction accuracy and risk-adjusted performance, offering a more robust tool for financial decision making in uncertain market environments.

## Keywords

LLM, Risk-Adjusted Return, Stock Price Prediction, Proximal Policy Optimization

## 1. Introduction

Stock price prediction is a critical task in financial markets, with significant implications for investors, analysts, and traders. Accurate forecasting can lead to better investment strategies, while inaccurate predictions can result in substantial fi-

nancial losses. Conventional stock price prediction techniques, including time series models, machine learning, and deep learning, typically concentrate on identifying price trends through historical data and technical indicators. For example, traditional machine learning models such as SVM, Random Forest, and Gradient Boosting are commonly used with technical indicators to predict stock prices [1]-[4]. However, due to the inherently non-linear nature of stock price movements, these models struggle to capture complex patterns effectively [5]. To address this limitation, researchers often turn to deep learning approaches, such as LSTM [6] and CNN [7], which are better suited for modeling intricate and non-linear dependencies in financial data.

To further enhance stock forecasting performance, researchers incorporate financial news sentiment alongside stock features and technical indicators [8]-[10]. These sentiments are derived from various sources, including tweets [11], news articles [12], and financial reports [13]. Before the emergence of Large Language Models (LLMs), sentiment analysis was typically conducted using traditional NLP tools such as VADER [14], which is a lexicon-based approach.

However, all these models and approaches typically overlook the inherent risks and uncertainties in the financial markets, leading to predictions that may be volatile or unreliable in the face of unforeseen market shifts.

Recent advancements in Natural Language Processing (NLP) have led to the creation of LLMs. These models have shown remarkable performance in various tasks, including sentiment analysis and the interpretation of financial news [10] [15] [16]. These models can process large amounts of textual data to extract valuable insights, making them promising tools for stock price prediction. Although LLMs can provide accurate predictions based on historical price data and market sentiment, they often do not explicitly account for the financial risks associated with their forecasts. This gap presents an opportunity to improve stock price prediction models by integrating risk considerations.

A prompt-based approach has recently gained attention as an effective method for leveraging LLMs in stock price prediction [15]-[17]. This prompt-based approach uses designed prompts to instruct the model in analyzing financial data, extracting key insights, and making informed predictions. By structuring the input with specific instructions, prompt-based LLMs can focus on relevant aspects of financial news, technical indicators, and market conditions, leading to more context-aware predictions. One major advantage of this approach is its adaptability, because prompts can be modified to reflect different market conditions, risk factors, or investor preferences without retraining the model [18]. Additionally, prompt engineering allows for the integration of multiple data sources, such as candlestick patterns and macroeconomic indicators, enabling a more comprehensive analysis. This makes prompt-based LLMs a flexible and efficient tool for stock forecasting, offering improved interpretability and alignment with financial decision making.

Moreover, recent studies have highlighted the potential of reinforcement learn-

ing in financial forecasting, particularly for portfolio optimization and risk management [19]–[21]. Reinforcement learning algorithms, such as PPO, have demonstrated success in dynamic environments by learning from past interactions and adjusting strategies accordingly [22]. However, limited research has explored the integration of LLM-generated predictions with reinforcement learning techniques to refine stock price forecasts in a risk-aware manner.

This research proposes a novel framework that integrates LLM-based stock price prediction with Proximal Policy Optimization (PPO), a reinforcement learning technique. The primary objective is to refine LLM predictions by incorporating risk-aware adjustments through PPO. By doing so, the framework addresses both the accuracy and the risk aspects of stock price forecasting. Specifically, we introduce a risk-adjusted reward function that takes into account financial risk metrics, such as Value at Risk (VaR) and Conditional Value at Risk (CVaR), to optimize the predictions. In finance, Value at Risk (VaR) is commonly used to estimate the potential loss of an asset or portfolio within a specified time period at a given confidence level. For example, a 95% VaR of \$1 million means there is a 95% chance that losses will not exceed \$1 million in the specified period. Conditional Value at Risk (CVaR), also referred to as Expected Shortfall, quantifies the average loss that exceeds the Value at Risk (VaR) threshold. It provides insight into extreme losses by averaging the worst-case scenarios that exceed VaR. CVaR is useful for assessing tail risk and is often preferred for risk management. By considering the risk-adjusted return, this framework aims to provide a more robust and reliable approach to stock price forecasting, ultimately offering a tool that better aligns with the complex dynamics of financial markets.

Through this approach, the paper contributes to both the application of LLMs in stock price prediction and the integration of risk-aware mechanisms in reinforcement learning models. The results show that the framework improves prediction accuracy while reducing financial risk. This makes it a useful tool for investors and analysts in a fast-changing financial market.

The main contributions of this research are as follows.

1. This research proposes a novel framework that combines LLMs with a risk-adjusted PPO to refine stock price predictions. The LLM generates initial price forecasts, while PPO adjusts these predictions based on financial risk, improving both prediction accuracy and stability.
2. The study introduces a risk-aware mechanism that integrates financial risk metrics (e.g., VaR and CVaR) into stock price prediction. This improves forecast reliability by considering market volatility, helping investors and analysts make better decisions.

The remainder of this paper is structured as follows. Firstly, related work is listed (Section 2). The methodology is then described (Section 3). In Section 4, the experimental results are presented. Finally, we conclude with Section 5.

## 2. Related Work

The task of stock price prediction has been extensively studied in financial and

computational research. Traditionally, stock price forecasting models rely on time-series analysis methods such as AutoRegressive Integrated Moving Average (ARIMA) and various forms of exponential smoothing [23]-[25]. However, these models are limited in their ability to capture complex non-linear relationships in financial markets. Therefore, machine learning and deep learning techniques have emerged as more effective tools for stock prediction due to their ability to model intricate patterns and trends.

### 2.1. Machine Learning for Stock Prediction

Machine learning models such as Support Vector Machines (SVMs) [26]-[28], Random Forests [5] [29] [30], and Gradient Boosting methods [31]-[33] have been used for predicting stock price movements. For example, Mahmoodi *et al.* [34] use a SVM model combined with metaheuristic algorithms to develop an advanced stock price forecasting model. Their approach integrates optimization techniques (metaheuristics) to enhance SVM's predictive accuracy, addressing challenges like non-linearity and volatility in financial markets. Vijh *et al.* [35] use a Random Forest model to predict the next day's close prices for five companies and results show the effectiveness of the model in stock price forecasting. However, these approaches with stock features and technical indicators as inputs have some limitations. For example, these approaches do not account for external factors like sentiment or macroeconomic events, which can significantly affect stock prices.

### 2.2. Deep Learning Approaches

Deep learning models, including Long Short Term Memory (LSTM) [36]-[38] and Convolutional Neural Network (CNN) [17] [39] [40], have gained attention in the stock prediction domain. LSTM models are particularly effective in modeling sequential data, such as time series, and have been shown to outperform traditional models in stock price prediction tasks [41]. On the other hand, CNNs have been used to process financial data in the form of images, such as candlestick charts, capturing spatial patterns that could be indicative of price movements [17] [42]. In order to further enhance the prediction performance, some hybrid models are proposed by some researchers. For example, Lu *et al.* [43] propose a CNN-LSTM model to forecast the stock price. CNN is used to extract features from the data. Then LSTM is fed with the extracted features to predict the stock price. Nikhil *et al.* [44] propose a Genetic Algorithm-Assisted LSTM-CNN (GA-LSTM-CNN) for stock price prediction. Although these models demonstrate improved performance over traditional methods, they often overlook the inherent risk associated with financial predictions, which makes them less robust under volatile market conditions.

### 2.3. Large Language Models for Stock Prediction

More recently, the use of LLMs for stock prediction has attracted attention due to

their ability to process large-scale textual data, such as financial news, analyst reports and social media sentiment. Models like BERT [45] and GPT [46] have been fine-tuned for financial sentiment analysis and have shown promising results in predicting market trends based on news articles and social media posts [47]. Even without financial fine-tuning, LLMs have the potential to outperform other traditional models in stock forecasting [17] [48]. These models, when combined with traditional stock data, offer a more comprehensive market view. They integrate both quantitative and qualitative factors. Recently, this is a growing trend of applying LLMs directly to forecast time series. For example, Wang *et al.* [49] propose a LLM called Stock Time for stock price forecasting and the results show that their proposed model can outperform recent LLMs with less memory usage and computational cost. However, the challenge remains in how to integrate LLM-based predictions with risk considerations to ensure financial stability.

#### 2.4. Reinforcement Learning for Stock Price Prediction

Reinforcement Learning (RL) has also been applied to stock price prediction, with the aim of optimizing trading strategies in a dynamic environment. For example, Altuner and Kilimci [50] use deep Q-learning with knowledge graph and community aware sentiments to predict the stock movement. Du and Shen [51] also use deep Q-learning with sentiment analysis to predict stocks in China. Furthermore, PPO, a popular RL algorithm [52], has been used to train agents that adjust portfolios or make buy/sell decisions based on market conditions [53]. Although a growing number of studies have shown the potential to use RL for stock trading by training agents to maximize returns while minimizing risk [54]-[56], most RL-based models focus primarily on optimizing returns, with less emphasis on accounting for financial risks, such as market volatility or drawdowns.

#### 2.5. Risk-Aware Prediction

Incorporating risk measures into stock prediction models is an emerging area of research. Traditional models that focus purely on prediction accuracy may lead to overly optimistic forecasts, which can result in significant losses during periods of high volatility. Risk-adjusted models, such as those incorporating Value at Risk (VaR) or Conditional Value at Risk (CVaR), offer a more robust framework by considering potential losses in the tail of the distribution [57].

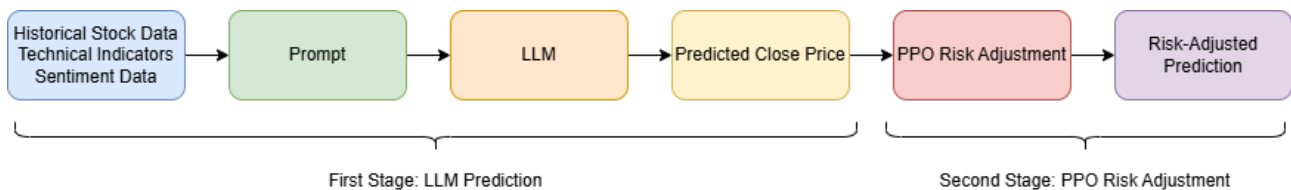
### 3. Methodology

This study proposes a two-stage framework (Figure 1) that integrates LLMs with a risk-aware reinforcement learning approach using PPO to improve stock price prediction. The first stage uses an LLM to generate initial stock price forecasts based on historical market data, technical indicators, and sentiment analysis. However, these predictions do not inherently account for risk, making them susceptible to market volatility. To address this, the second stage introduces a PPO-based adjustment mechanism that refines the LLM forecasts by incorporating

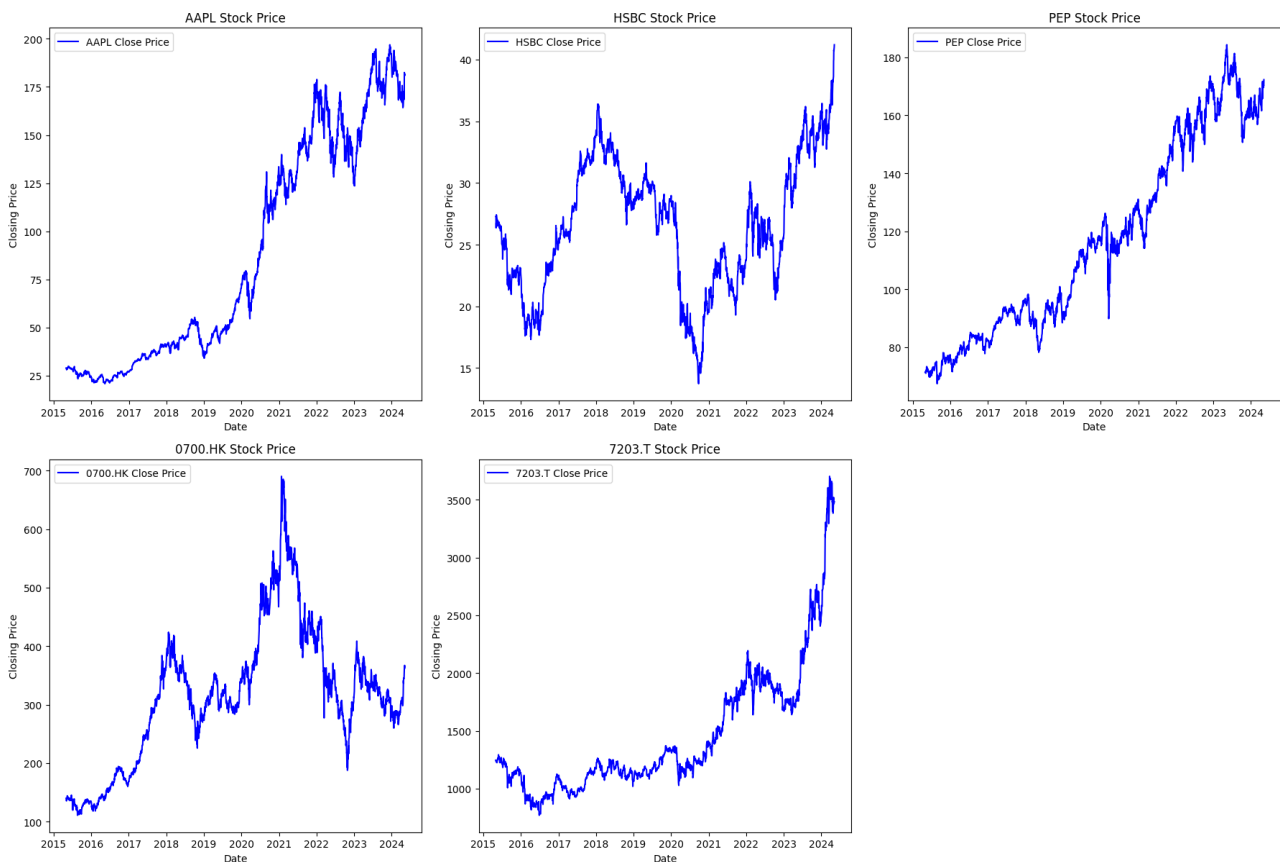
risk-sensitive modifications. This framework aims to strike a balance between prediction accuracy and financial stability, offering a more robust approach to stock price forecasting.

### 3.1. Data

The data used in this study consist of historical stock price data, technical indicators, and sentiment analysis. Stock price data, including close price, open price, high, low, and trading volume, are collected for a set of companies from 2015 to 2024 by using the Python library yfinance. These companies include Apple, HSBC, Pepsi, Tencent (0700.HK), and Toyota (7203.T). The daily close price trends for these five companies are illustrated in **Figure 2**. The summary statistics of stock prices are shown in **Table 1**. The data set is prepared with a sliding window of five days, where the most recent five days of data are used to predict the



**Figure 1.** Proposed two-stage framework for stock forecasting



**Figure 2.** Daily close price of five companies.

**Table 1.** Summary statistics of stock prices.

	Apple	HSBC	Pepsi	Tencent	Toyota
Count	2270	2270	2270	2167	2147
Mean	87.36	26.50	116.81	314.49	1457.08
Std Dev	58.04	5.27	33.08	117.52	553.96
Min	20.67	13.73	67.36	110.97	768.73
th Percentile	35.90	22.58	88.54	244.21	1112.45
Median	58.75	26.58	113.80	312.70	1224.96
th Percentile	144.71	30.60	151.23	370.07	1776.84
Max	196.93	41.20	184.35	690.13	3703.98

**Table 2.** Descriptions of technical indicators.

Indicator	Description
SMA (Simple Moving Average)	The average price over a set period, giving equal weight to all prices in the period.
EMA (Exponential Moving Average)	A moving average that gives more weight to recent prices, making it more responsive to new data.
RSI (Relative Strength Index)	Measures price movement speed and change; identifies overbought or oversold conditions.
MACD (Moving Average Convergence Divergence)	Tracks the difference between short-term and long-term moving averages to identify trends.
Bollinger Bands	A volatility indicator consisting of a moving average with upper and lower bands based on standard deviations.

**Table 3.** LLM Prompt.

Prompt Description
You are a financial analyst with expertise in stock market forecasting.
Your task is to analyze market data and predict the next trading day stock price.
Use historical price trends, technical indicators, and sentiment analysis to provide an informed forecast.
Ensure that your predictions are well-justified, considering multiple financial factors.
. <b>Predicted Stock Price:</b> The forecasted close price for the next trading day.
. <b>Price Movement Likelihood:</b> The likelihood of the predicted stock price.
. <b>Justification:</b> Provide an explanation for the predicted stock price and the corresponding likelihood, considering the following:
-Historical market data (e.g., recent closing prices).
-Technical indicators (e.g., SMA, EMA, RSI, MACD, Bollinger Bands).
-Sentiment analysis (e.g., news sentiment, market sentiment).
Please weigh these signals and justify the predicted stock price.

close price of the next day. In addition, technical indicators such as Simple Moving Average (SMA), Exponential Moving Average (EMA), Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands



are derived from the raw price data by using the Python library pandas-ta. A brief description of each technical indicator is given in **Table 2**. Technical indicators are calculated from historical prices and volumes and they are often used by traders and investors to predict the stock movement in finance. Sentiment analysis is conducted on financial news using polarity scores, which help capture market sentiment that could influence stock price movements. The Python package NLTK is used to generate the polarity scores of financial news. The polarity score measures the sentiment of text and it ranges from  $-1$  to  $1$ . A score closer to  $1$  indicates positive sentiment, while a score near  $-1$  represents negative sentiment. A score around  $0$  suggests neutral sentiment. Daily news articles are downloaded from reliable financial news sources such as Yahoo Finance using the Financial News Feed and Stock News Sentiment data API<sup>1</sup>. The first 70% of the data is used for training and the remaining 30% of the data is used for testing.

## 3.2. Model Setup

### 3.2.1. LLM-Based Stock Price Prediction

The first stage of the framework involves using an LLM to predict stock prices. The LLM receives structured prompt input that includes historical market data, technical indicators, and sentiment analysis. This information is processed to generate a forecast for the next day's stock price. The LLM model used in this study is "deepseek-r1-distill-llama-70b". The LLM is prompted to generate stock price predictions, providing the predicted price, its corresponding likelihood, and explanations for these forecasts based on the input features. **Table 3** shows the structure and content of the prompt used in this study. Although the LLM produces a valuable initial prediction, it does not directly account for financial risks or market volatility, which requires the integration of the PPO model in the second stage.

### 3.2.2. Risk-Aware PPO Adjustment

The second stage of the framework introduces a Proximal Policy Optimization (PPO) agent to adjust the LLM predictions by incorporating risk-sensitive modifications. PPO, proposed by [52], is selected due to its stability in policy optimization and its ability to ensure that drastic policy changes are avoided. The PPO agent operates within an environment where it takes the LLM-predicted stock price and refines it based on a defined reward function that accounts for both prediction accuracy and financial risk. The state representation (Equation 1) for the PPO agent includes the LLM-predicted price, historical price trends, market volatility, and risk metric, VaR. The PPO agent's action space (Equation 2) allows it to adjust the predicted price, ensuring that any forecast modifications stay within a reasonable range to prevent large fluctuations.

The reward function (Equation 4) in PPO is designed to balance predictive accuracy and risk control. It is formulated as a weighted combination of the negative absolute prediction error and a risk penalty term that incorporates risk measure CVaR. PPO is trained using historical stock data, where it learns the optimal ad-

<sup>1</sup><https://eodhd.com/financial-apis/stock-market-financial-news-api>.



justment policy by interacting with the market environment and receiving feedback based on its actions. Training involves updating the PPO model's policy network to maximize the cumulative reward while maintaining stable updates through a clipped surrogate objective function (Equation 5).

$$s_t = \{\hat{y}_t, P_t, \sigma_t, \text{VaR}_t\} \quad (1)$$

- $s_t$  represents the state at time  $t$ .
- $\hat{y}_t$  is the LLM-predicted stock price.
- $P_t$  is the historical price trend.
- $\sigma_t$  is the market volatility.
- $\text{VaR}_t$  is the Value at Risk.

$$y_t = \hat{y}_t + a_t \quad (2)$$

- $y_t$  is the adjusted stock price prediction.
- $a_t$  is the adjustment made by the PPO agent.

$$|a_t| \leq \epsilon \cdot \hat{y}_t \quad (3)$$

- $\epsilon$  is a predefined constraint ensuring the adjustment remains within a reasonable range, preventing large fluctuations.

$$R_t = -|\hat{y}_t - y_t^*| - \lambda \cdot \text{CVaR}_\alpha \quad (4)$$

- $R_t$  is the reward at time  $t$ .
- $y_t^*$  is the true stock price.
- $\lambda$  is a weighting factor.
- $\text{CVaR}_\alpha$  represents the Conditional Value at Risk at confidence level  $\alpha$ .

$$L(\theta) = \mathbb{E}_t \left[ \min \left( r_t(\theta) A_t, \text{clip} \left( r_t(\theta), 1 - \epsilon, 1 + \epsilon \right) A_t \right) \right] \quad (5)$$

- $L(\theta)$  is the PPO objective function.
- $r_t(\theta)$  is the probability ratio between the new and old policies.
- $A_t$  is the advantage estimate.
- $\epsilon$  is a clipping parameter to stabilize training.

$$r_t(\theta) = \frac{\pi_\theta(a_t | s_t)}{\pi_{\theta_{\text{old}}}(a_t | s_t)} \quad (6)$$

- $r_t(\theta)$  is the probability ratio of taking action  $a_t$  under the new policy  $\pi_\theta$  versus the old policy  $\pi_{\theta_{\text{old}}}$ .

### 3.3. Evaluation

The performance of the proposed model is evaluated by comparing the adjusted predictions from the LLM-PPO framework with predictions from other models. These models include a baseline LLM model that makes predictions without any adjustments, traditional machine learning models like SVR, XGBoost, TCN and LSTM, and the proposed LLM-PPO hybrid model. The evaluation criteria consist of multiple aspects to assess both prediction accuracy and risk-adjusted returns.

Prediction accuracy is measured using Mean Absolute Percentage Error (MAPE)

(Equation 7), where the predicted stock prices are compared with the actual stock prices, and Root Mean Square Error (RMSE) (Equation 8), which quantifies the average magnitude of errors by taking the square root of the average squared differences between predicted and actual values. To evaluate the financial stability of the predictions, risk-adjusted return metrics such as the Sharpe ratio and Sortino ratio are used. The Sharpe ratio (Equation 9) evaluates the return earned for each unit of risk taken, whereas the Sortino ratio (Equation 10) focuses specifically on downside risk. Additionally, the maximum drawdown (Equation 11) is calculated to assess the largest peak-to-trough loss during the evaluation period.

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right| \times 100 \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (A_i - F_i)^2} \quad (8)$$

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p} \quad (9)$$

$$\text{Sortino Ratio} = \frac{R_p - R_f}{\sigma_d} \quad (10)$$

$$\text{Max Drawdown} = \max \left( \frac{\text{Peak} - \text{Trough}}{\text{Peak}} \right) \quad (11)$$

Where:

- $A_i$  is the actual value, and  $F_i$  is the forecasted value.
- $R_p$  is the stock return,  $R_f$  is the risk-free rate.
- $\sigma_p$  is the standard deviation of returns, and  $\sigma_d$  is the standard deviation of downside deviations.
- Peak and Trough refer to the highest and lowest values of the stock over a given period.

### 3.3.1. Baseline Models

Below is a brief overview of each baseline model utilized in this study.

**SVR (Support Vector Regression):** SVR is a regression technique that applies the principles of Support Vector Machines (SVM) to predict continuous values. It aims to find a function that best fits the data by mapping it into a higher-dimensional space and minimizing the prediction error within a margin.

**XGBoost (Extreme Gradient Boosting):** XGBoost is a robust machine learning algorithm that utilizes gradient boosting. It constructs an ensemble of decision trees, with each new tree aiming to rectify the mistakes of its predecessors.

**TCN (Temporal Convolutional Network):** TCN is a deep learning architecture designed for sequence modeling tasks. It employs convolutional layers with causal padding to handle sequential data and capture long-term dependencies. Unlike recurrent networks, TCNs can efficiently process long sequences due to their parallelization capabilities, making them suitable for time series forecasting tasks like stock price prediction.

**LSTM (Long Short Term Memory):** LSTM is a type of recurrent neural network (RNN) designed to address the vanishing gradient problem in traditional RNNs. It uses specialized gates to regulate the flow of information and maintain long-term dependencies in sequential data. LSTMs are widely used in time series analysis and natural language processing tasks due to their ability to learn from historical data over extended periods.

## 4. Experimental Setup

The LLM used for this study is based on the “deepseek-r1-distill-llama-70b” model, which is from the American AI company Groq AI<sup>2</sup>. Users can use this model with API but fine-tuning is not available. However, this model has been shown to be capable of generating competitive performance in the previous study even without fine-tuning [17]. The “deepseek-r1-distill-llama-70B” is a distilled version of the Llama 70B model. It is designed to be smaller and more efficient while maintaining strong performance. The model is optimized for tasks like natural language understanding, text generation, and reasoning. Due to the distillation process, it runs faster and requires less computing power compared to the original 70B model.

The PPO algorithm is implemented using Stable-Baselines3 in PyTorch, leveraging its ability to handle reinforcement learning tasks efficiently. The training process involves updating the PPO model using historical stock data, where the agent learns to adjust stock price predictions based on the reward function. The PPO model is trained for 100000 timesteps with a learning rate of  $3e-4$ , a clip range of 0.2, and a discount factor (gamma) of 0.99. The entropy coefficient is set to 0.01 to balance exploration and exploitation.

## 5. Results

The experiment results, shown in **Table 4**, compare the performance of different models across five stocks. Six key metrics are presented: MAPE (**Figure 3**), RMSE (**Figure 4**), Sharpe Ratio, Sortino Ratio, Maximum Drawdown, and Cumulative Return (**Figure 5**). Lower MAPE and RMSE values indicate better prediction accuracy, while a lower Maximum Drawdown reflects better risk management. Higher Sharpe and Sortino Ratios suggest better risk-adjusted returns, and Cumulative Return shows the overall profitability of each model.

The results show that LLM-PPO performs the best across all five stocks in both prediction accuracy and risk management. For Apple, LLM-PPO not only has the lowest prediction errors but also generates the highest cumulative returns and the smallest drawdown. The same trend is observed with Pepsi, where LLM-PPO reinforcement learning framework effectively handles price fluctuations, providing both accurate predictions and strong risk-adjusted returns.

In the case of HSBC, XGBoost and LSTM perform well in terms of prediction accuracy, but LLM-PPO still outperforms in overall performance. Traditional

<sup>2</sup><https://groq.com/>.

**Table 4.** Comparison of model performance across multiple stocks.

Stock	Model	MAPE (%) ↓	RMSE ↓	Sharpe Ratio ↑	Sortino Ratio ↑	Max Drawdown ↓	Cumulative Return ↑
*Apple	LLM (Baseline)	0.1325	0.0578	0.9203	0.0049	−0.3802	0.1750
	XGBoost	0.0991	0.0447	0.7801	0.0032	−0.2955	0.1853
	SVR	0.1173	0.0731	0.7105	0.0028	−0.3105	0.1482
	TCN	0.1019	0.0603	0.8452	0.0031	−0.3422	0.1723
	LSTM	0.1489	0.1048	0.6310	0.0024	−0.3899	0.1125
	LLM-PPO	<b>0.0850</b>	<b>0.0366</b>	<b>1.0213</b>	0.0035	<b>−0.2799</b>	<b>0.2321</b>
*HSBC	LLM (Baseline)	0.1278	0.0469	0.5303	0.0018	−0.6215	0.1285
	XGBoost	0.0952	0.0203	0.5402	0.0019	−0.2520	0.1367
	SVR	0.1104	0.0483	0.4217	0.0015	−0.3511	0.1083
	TCN	0.1048	0.0379	0.4794	0.0020	−0.2993	0.1295
	LSTM	0.1019	0.0372	0.3789	0.0017	−0.4014	0.1197
	LLM-PPO	<b>0.0812</b>	<b>0.0186</b>	<b>1.0337</b>	<b>0.0022</b>	<b>−0.2215</b>	<b>0.2126</b>
*Pepsi	LLM (Baseline)	0.1407	0.0539	0.4485	0.0022	−0.2897	0.1391
	XGBoost	0.0944	0.0382	0.5963	0.0024	−0.2480	0.1582
	SVR	0.1095	0.0451	0.5018	0.0023	−0.2979	0.1480
	TCN	0.1179	0.0838	0.5496	0.0027	−0.3476	0.1302
	LSTM	0.1246	0.0785	0.5750	0.0025	−0.3680	0.1109
	LLM-PPO	<b>0.0786</b>	<b>0.0322</b>	<b>1.0427</b>	<b>0.0030</b>	<b>−0.1989</b>	<b>0.2203</b>
*Tencent	LLM (Baseline)	0.1368	0.0487	0.4683	0.0023	−0.3487	0.1312
	XGBoost	0.0918	0.0336	0.6176	0.0025	−0.2781	0.1684
	SVR	0.1089	0.0439	0.5483	0.0022	−0.3297	0.1492
	TCN	0.1149	0.0764	0.5202	0.0028	−0.3189	0.1415
	LSTM	0.1273	0.1064	0.5982	0.0026	−0.3465	0.1194
	LLM-PPO	<b>0.0763</b>	<b>0.0295</b>	<b>1.0519</b>	<b>0.0031</b>	<b>−0.1799</b>	<b>0.2396</b>
*Toyota	LLM (Baseline)	0.1448	0.0558	0.4957	0.0025	−0.4160	0.1296
	XGBoost	0.1092	0.0227	0.5989	0.0027	−0.2773	0.1612
	SVR	0.2499	0.0347	0.4972	0.0024	−0.3312	0.1388
	TCN	0.2146	0.0412	0.5463	0.0026	−0.3097	0.1490
	LSTM	0.2309	0.0354	0.4691	0.0022	−0.3462	0.1213
	LLM-PPO	<b>0.0867</b>	<b>0.0189</b>	<b>1.0634</b>	<b>0.0032</b>	<b>−0.1825</b>	<b>0.2239</b>

models like SVR and TCN show larger drawdowns and weaker risk-adjusted returns. This highlights LLM-PPO's strength, even in stable markets, where it minimizes losses while capturing gains effectively.

For Toyota, LSTM shows decent forecasting accuracy, but its risk management is not as strong as LLM-PPO. The baseline LLM model performs well, but the PPO-enhanced version significantly improves all metrics, especially in reducing

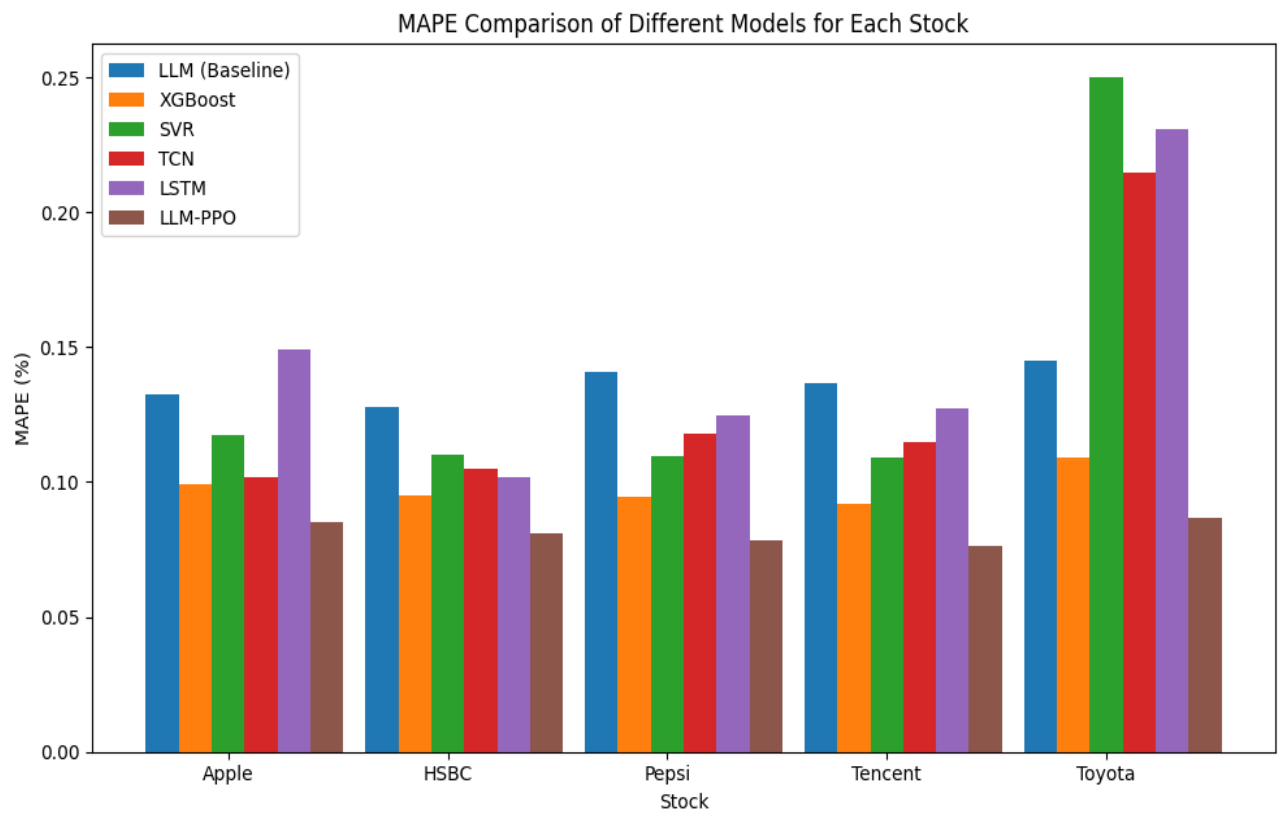


Figure 3. MAPEs of All models for each stock.

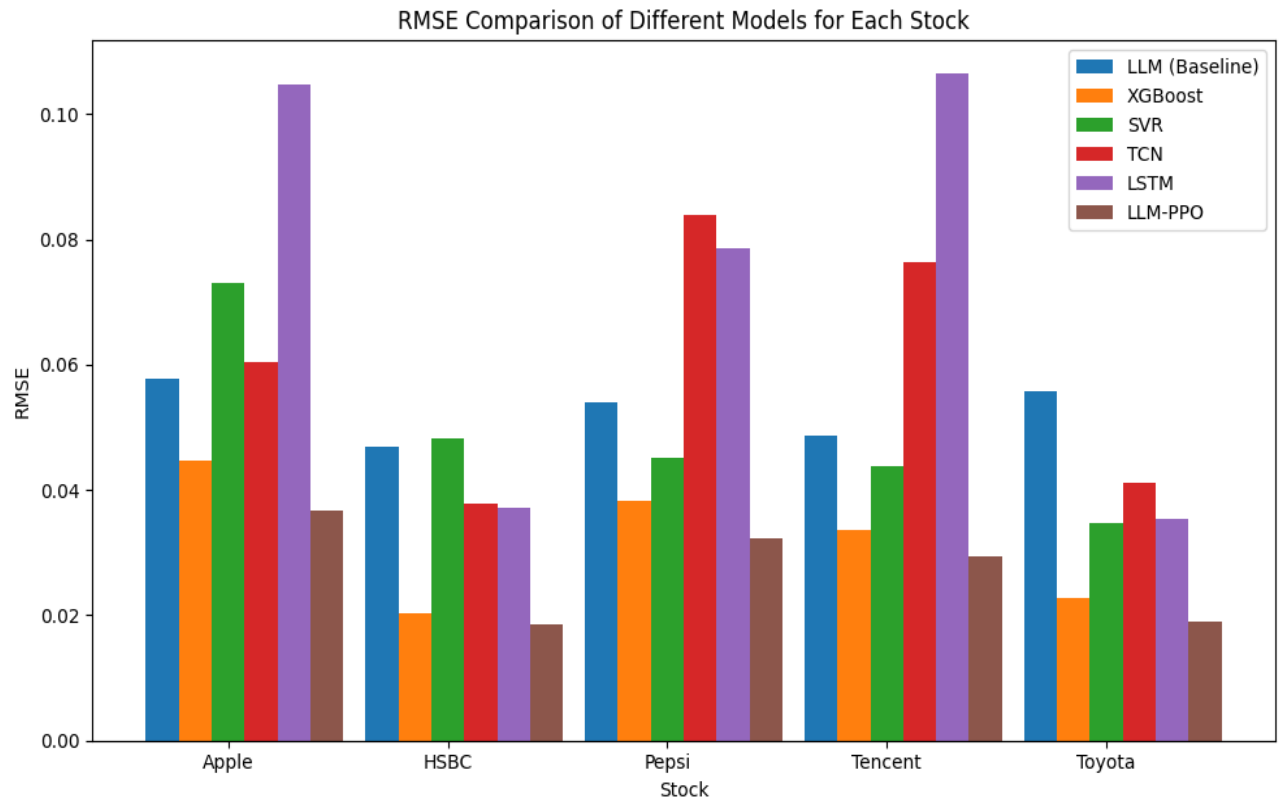
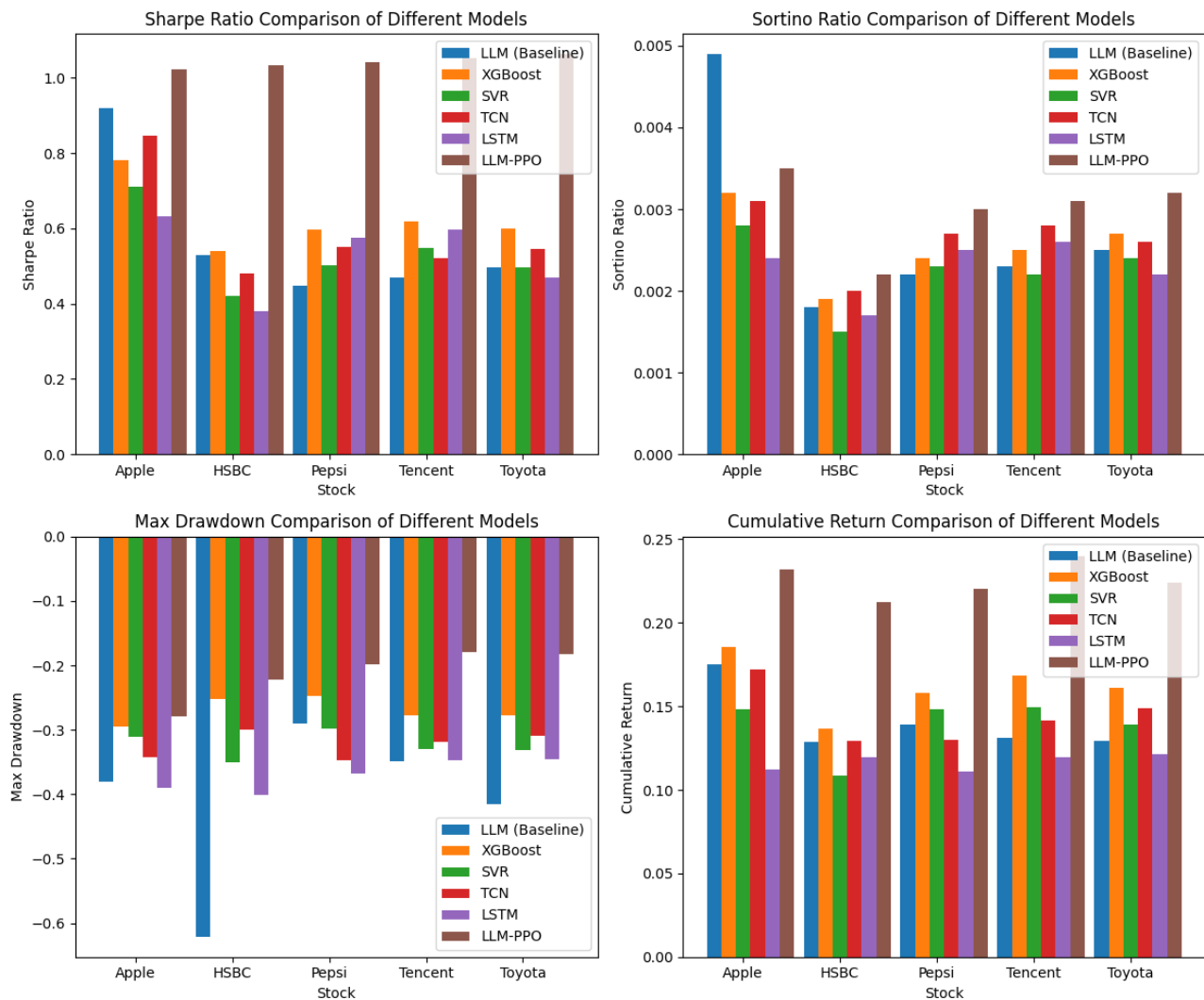


Figure 4. RMSEs of All models for each stock.



**Figure 5.** Econometric metrics across different models for stock price prediction.

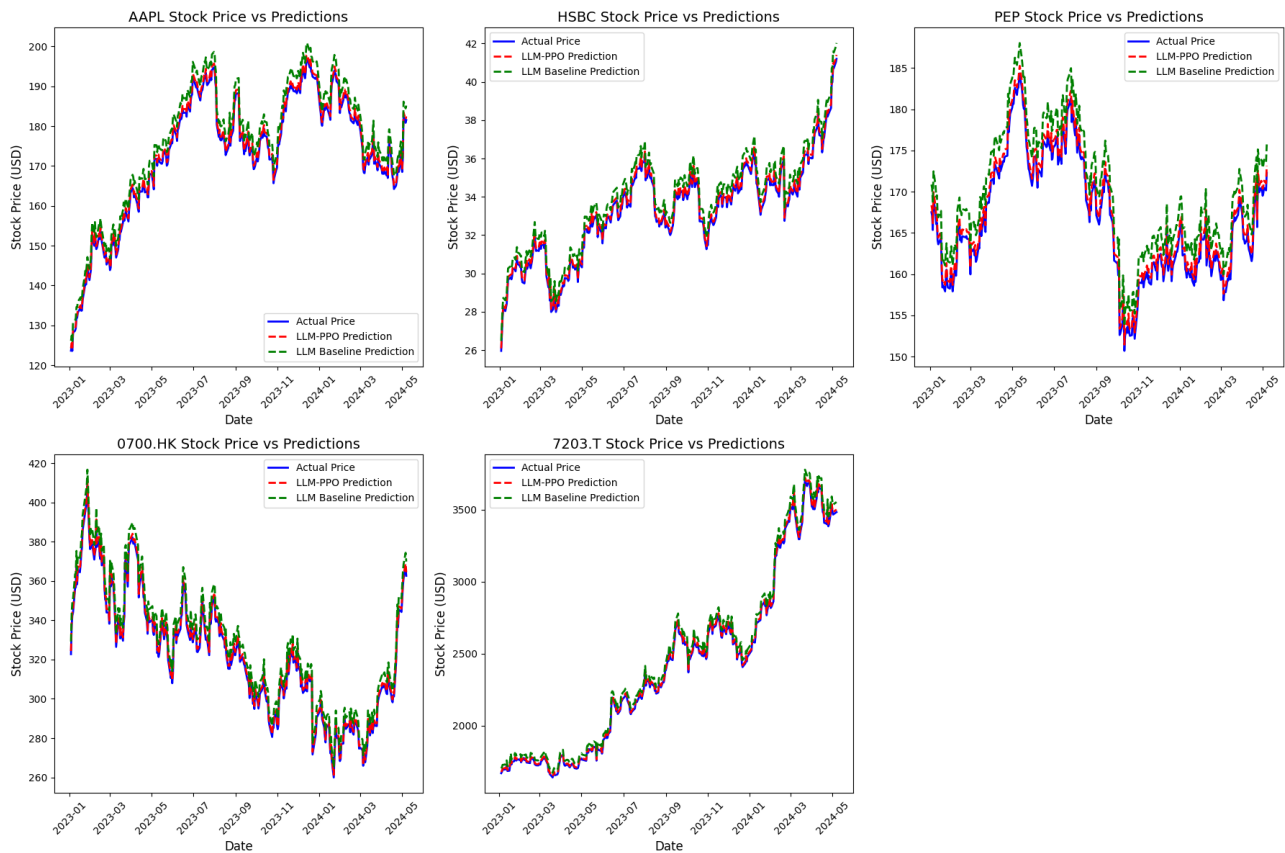
drawdowns and increasing cumulative returns. This demonstrates the value of using reinforcement learning to enhance predictions.

**Figure 6** visualizes the prediction performance of both the baseline LLM and the LLM-PPO models. It is clear that the LLM-PPO model more effectively captures stock movements compared to the baseline LLM without any enhancements or adjustments.

According to **Figure 3** and **Figure 4**, another important finding is that although LLM without RL adjustment cannot outperform traditional machine learning models, it can generate competitive results compared to the traditional models, which indicates LLM's potential in time series forecasting.

Overall, LLM-PPO provides the best balance of low prediction errors, strong risk-adjusted returns, and controlled drawdowns across all stocks. It reduces maximum drawdowns by 20% - 40% compared to other models. These results suggest that LLM-PPO combination of predictive modeling and risk-sensitive reinforcement learning creates a more reliable and effective solution for stock forecasting,

excelling not only in accuracy but also in providing real-world investment value.



**Figure 6.** Prediction and actual stock price example.

## 6. Conclusions

This methodology presents a novel approach that combines LLM-based stock forecasting with risk-aware reinforcement learning. By integrating PPO, the proposed framework ensures that stock price predictions are not only accurate but also aligned with financial risk constraints. This hybrid approach has the potential to improve decision making in financial markets by providing more stable and reliable predictions, which can be beneficial for investors and financial analysts navigating market uncertainties.

However, there are several limitations to this approach that warrant further research. For example, the computational cost of training LLMs and PPO agents is high, which may restrict scalability in real-world trading applications. Future studies could explore more efficient architectures or alternative reinforcement learning methods to reduce training complexity. In addition, although the framework improves stock price prediction and risk management, its effectiveness in different market conditions, such as low-liquidity stocks or extreme volatility periods, remains an open question. Expanding the analysis to a broader set of assets and conducting real-world trading simulations would help validate its generalizability and practical applicability.



## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Lin, Y., Guo, H. and Hu, J. (2013) An SVM-Based Approach for Stock Market Trend Prediction. *The 2013 International Joint Conference on Neural Networks (IJCNN)*, Dallas, 4-9 August 2013, 1-7. <https://doi.org/10.1109/ijcnn.2013.6706743>
- [2] Fenghua, W., Jihong, X., Zhifang, H. and Xu, G. (2014) Stock Price Prediction Based on SSA and SVM. *Procedia Computer Science*, **31**, 625-631. <https://doi.org/10.1016/j.procs.2014.05.309>
- [3] Omar, A.B., Huang, S., Salameh, A.A., Khurram, H. and Fareed, M. (2022) Stock Market Forecasting Using the Random Forest and Deep Neural Network Models before and during the COVID-19 Period. *Frontiers in Environmental Science*, **10**, Article 917047. <https://doi.org/10.3389/fenvs.2022.917047>
- [4] Zeng, X., Wang, F., Mai, X. and Cai, J. (2024) Stock Price Prediction Model Based on I-NSGAI-XGBoost Algorithm. *Proceedings of the 2024 4th International Conference on Artificial Intelligence, Big Data and Algorithms*, Zhengzhou, 21-23 June 2024, 800-805. <https://doi.org/10.1145/3690407.3690541>
- [5] Chen, Q. and Kawashima, H. (2024) Stock Price Prediction Using LLM-Based Sentiment Analysis. *2024 IEEE International Conference on Big Data (Big Data)*, Washington, 15-18 December 2024, 4846-4853. <https://doi.org/10.1109/bigdata62323.2024.10825946>
- [6] Shinde, S., Wadhwa, L., Mohane, N., Pagar, V., Sherje, N. and Mane, S. (2023) Stock Price Prediction Using LSTM. *2023 7th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Pune, 18-19 August 2023, 1-7. <https://doi.org/10.1109/iccubea58933.2023.10392023>
- [7] Wojarnik, G. (2023) The Potential of Convolutional Neural Networks for the Analysis of Stock Charts. *Procedia Computer Science*, **225**, 941-950. <https://doi.org/10.1016/j.procs.2023.10.081>
- [8] Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P. and Anastasiu, D.C. (2019) Stock Price Prediction Using News Sentiment Analysis. *2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (Big Data Service)*, Newark, 4-9 April 2019, 205-208. <https://doi.org/10.1109/bigdataservice.2019.00035>
- [9] Maqbool, J., Aggarwal, P., Kaur, R., Mittal, A. and Ganaie, I.A. (2023) Stock Prediction by Integrating Sentiment Scores of Financial News and MLP-Regressor: A Machine Learning Approach. *Procedia Computer Science*, **218**, 1067-1078. <https://doi.org/10.1016/j.procs.2023.01.086>
- [10] Gu, W.J., Zhong, Y.H., Li, S.Z., Wei, C.S., *et al.* (2024) Predicting Stock Prices with Finbert-LSTM: Integrating News Sentiment Analysis. *Proceedings of the 2024 8th International Conference on Cloud and Big Data Computing*, Oxford, 15-17 August 2024, 67-72. <https://doi.org/10.1145/3694860.3694870>
- [11] Sariyer, M., Akil, A., Bulgurcu, F.N., Oge, F.E. and Ganiz, M.C. (2022) Individual Stock Price Prediction by Using KAP and Twitter Sentiments with Machine Learning for BIST30. *2022 International Conference on Innovations in Intelligent Systems and Applications (INISTA)*, Biarritz, 8-12 August 2022, 1-6. <https://doi.org/10.1109/inista55318.2022.9894172>
- [12] Kaya, M.I.Y. and Karsligil, M.E. (2010) Stock Price Prediction Using Financial News Articles. *2010 2nd IEEE International Conference on Information and Financial En-*

- gineering*, Chongqing, 17-19 September 2010, 478-482.  
<https://doi.org/10.1109/icife.2010.5609404>
- [13] Faccia, A., McDonald, J. and George, B. (2023) NLP Sentiment Analysis and Accounting Transparency: A New Era of Financial Record Keeping. *Computers*, **13**, Article 5. <https://doi.org/10.3390/computers13010005>
  - [14] Ekaputri, A.P. and Akbar, S. (2022) Financial News Sentiment Analysis Using Modified VADER for Stock Price Prediction. 2022 9th *International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, Tokoname, 28-29 September 2022, 1-6. <https://doi.org/10.1109/icaicta56449.2022.9932925>
  - [15] Xie, Q.Q., Han, W.G., Zhang, X., Lai, Y.Z., *et al.* (2023) PIXIU: A Large Language Model, Instruction Data and Evaluation Benchmark for Finance.
  - [16] Yang, H., Liu, X. and Dan Wang, C. (2023) FinGPT: Open-Source Financial Large Language Models. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4489826>
  - [17] Chen, Q.Z. (2025) Stock Price Change Prediction Using Prompt-Based LLMs with RL-Enhanced Post-Hoc Adjustments. *Proceedings of the 4th International Conference on Bigdata Blockchain and Economy Management*, Wuhan, 14-16 March 2025.
  - [18] Xue, S.Q., Zhou, F., Xu, Y., *et al.* (2024) WeaverBird: Empowering Financial Decision-Making with Large Language Model, Knowledge Base, and Search Engine.
  - [19] Kabbani, T. and Duman, E. (2022) Deep Reinforcement Learning Approach for Trading Automation in the Stock Market. *IEEE Access*, **10**, 93564-93574.  
<https://doi.org/10.1109/access.2022.3203697>
  - [20] Li, Y., Liu, P. and Wang, Z. (2022) Stock Trading Strategies Based on Deep Reinforcement Learning. *Scientific Programming*, **2022**, 1-15.  
<https://doi.org/10.1155/2022/4698656>
  - [21] Fiorini, P.M. and Fiorini, P. (2021) A Simple Reinforcement Learning Algorithm for Stock Trading. 2021 11th *IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, Cracow, 22-25 September 2021, 824-830.  
<https://doi.org/10.1109/idaacs53288.2021.9660900>
  - [22] Zou, J., Lou, J., Wang, B. and Liu, S. (2024) A Novel Deep Reinforcement Learning Based Automated Stock Trading System Using Cascaded LSTM Networks. *Expert Systems with Applications*, **242**, Article 122801.  
<https://doi.org/10.1016/j.eswa.2023.122801>
  - [23] Ariyo, A.A., Adewumi, A.O. and Ayo, C.K. (2014) Stock Price Prediction Using the ARIMA Model. 2014 *UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, Cambridge, 26-28 March 2014, 106-112.  
<https://doi.org/10.1109/uksim.2014.67>
  - [24] Alotaibi, R. (2022) ARIMA Model for Stock Market Prediction. 2022 8th *International Conference on Computer Technology Applications*, Vienna, 12-14 May 2022, 1-4. <https://doi.org/10.1145/3543712.3543723>
  - [25] Chatterjee, A., Bhowmick, H. and Sen, J. (2021) Stock Price Prediction Using Time Series, Econometric, Machine Learning, and Deep Learning Models. 2021 *IEEE Mysore Sub Section International Conference (MysuruCon)*, Hassan, 24-25 October 2021, 289-296. <https://doi.org/10.1109/mysurucon52639.2021.9641610>
  - [26] Ricky, A., Genesius, K., Setiawan, M., Nugroho Jaya, R. and Suhartono, D. (2022) Stock Market Prediction Using Support Vector Regressor with Radial Basis Function Kernel. 2022 *the 5th International Conference on Image and Graphics Processing (ICIGP)*, Beijing, 7-9 January 2022, 380-387.  
<https://doi.org/10.1145/3512388.3512444>

- [27] Chhajaj, P., Shah, M. and Kshirsagar, A. (2022) The Applications of Artificial Neural Networks, Support Vector Machines, and Long-Short Term Memory for Stock Market Prediction. *Decision Analytics Journal*, **2**, Article 100015. <https://doi.org/10.1016/j.dajour.2021.100015>
- [28] Cao, R., Liang, X. and Ni, Z. (2012) Stock Price Forecasting with Support Vector Machines Based on Web Financial Information Sentiment Analysis. In: *Lecture Notes in Computer Science*, Springer, 527-538. [https://doi.org/10.1007/978-3-642-35527-1\\_44](https://doi.org/10.1007/978-3-642-35527-1_44)
- [29] Du, S., Hao, D. and Li, X. (2022) Research on Stock Forecasting Based on Random Forest. 2022 *IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA)*, Dalian, 28-30 October 2022, 301-305. <https://doi.org/10.1109/icdsca56264.2022.9987903>
- [30] Talazadeh, S. and Peraković, D. (2024) SARF: Stock Market Prediction with Sentiment-Augmented Random Forest. In: *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Springer, 54-64. [https://doi.org/10.1007/978-3-031-72393-3\\_5](https://doi.org/10.1007/978-3-031-72393-3_5)
- [31] Gumelar, A.B., Setyorini, H., Adi, D.P., Nilowardono, S., Latipah, W.A., *et al.* (2020) Boosting the Accuracy of Stock Market Prediction Using XGBoost and Long Short-Term Memory. 2020 *International Seminar on Application for Technology of Information and Communication (iSemantic)*, Semarang, 19-20 September 2020, 609-613. <https://doi.org/10.1109/isemantic50169.2020.9234256>
- [32] Hu, Y., Shao, L., La, L. and Hua, H. (2021) Using Investor and News Sentiment in Tourism Stock Price Prediction Based on XGBoost Model. 2021 *IEEE/ACIS 6th International Conference on Big Data, Cloud Computing, and Data Science (BCD)*, Zhuhai, 13-15 September 2021, 20-24. <https://doi.org/10.1109/bcd51206.2021.9581619>
- [33] Deng, S., Huang, X., Zhu, Y., Su, Z., Fu, Z. and Shimada, T. (2023) Stock Index Direction Forecasting Using an Explainable Extreme Gradient Boosting and Investor Sentiments. *The North American Journal of Economics and Finance*, **64**, Article 101848. <https://doi.org/10.1016/j.najef.2022.101848>
- [34] Mahmoodi, A., Hashemi, L., Jasemi, M., Mehraban, S., Laliberté, J. and Millar, R.C. (2022) A Developed Stock Price Forecasting Model Using Support Vector Machine Combined with Metaheuristic Algorithms. *American Journal of Open Research*, **60**, 59-86. <https://doi.org/10.1007/s12597-022-00608-x>
- [35] Vijh, M., Chandola, D., Tikkiwal, V.A. and Kumar, A. (2020) Stock Closing Price Prediction Using Machine Learning Techniques. *Procedia Computer Science*, **167**, 599-606. <https://doi.org/10.1016/j.procs.2020.03.326>
- [36] Sang, S. and Li, L. (2024) A Novel Variant of LSTM Stock Prediction Method Incorporating Attention Mechanism. *Mathematics*, **12**, Article 945. <https://doi.org/10.3390/math12070945>
- [37] Md, A.Q., Kapoor, S., A.V., C.J., Sivaraman, A.K., Tee, K.F., H., S., *et al.* (2023) Novel Optimization Approach for Stock Price Forecasting Using Multi-Layered Sequential LSTM. *Applied Soft Computing*, **134**, Article 109830. <https://doi.org/10.1016/j.asoc.2022.109830>
- [38] Nadif, M.A., Rahman Samin, M.T. and Islam, T. (2022) Stock Market Prediction Using Long Short-Term Memory (LSTM). 2022 *Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, Bhilai, 21-22 April 2022, 1-6. <https://doi.org/10.1109/icaect54875.2022.9807655>
- [39] Sayavong, L., Wu, Z. and Chalita, S. (2019) Research on Stock Price Prediction

- Method Based on Convolutional Neural Network. 2019 *International Conference on Virtual Reality and Intelligent Systems (ICVRIS)*, Jishou, 14-15 September 2019, 173-176. <https://doi.org/10.1109/icvr.2019.00050>
- [40] Mehtab, S. and Sen, J. (2020) Stock Price Prediction Using CNN and LSTM-Based Deep Learning Models. 2020 *International Conference on Decision Aid Sciences and Application (DASA)*, Sakheer, 8-9 November 2020, 447-453. <https://doi.org/10.1109/dasa51403.2020.9317207>
- [41] Gülmez, B. (2023) Stock Price Prediction with Optimized Deep LSTM Network with Artificial Rabbits Optimization Algorithm. *Expert Systems with Applications*, **227**, Article 120346. <https://doi.org/10.1016/j.eswa.2023.120346>
- [42] Zhou, L., Zhang, Y., Yu, J., Wang, G., Liu, Z., Yongchareon, S., *et al.* (2025) LLM-Augmented Linear Transformer-CNN for Enhanced Stock Price Prediction. *Mathematics*, **13**, Article 487. <https://doi.org/10.3390/math13030487>
- [43] Lu, W., Li, J., Li, Y., Sun, A. and Wang, J. (2020) A CNN-LSTM-Based Model to Forecast Stock Prices. *Complexity*, **2020**, 1-10. <https://doi.org/10.1155/2020/6622927>
- [44] Nikhil, S., Sah, R.K., Kumar Parki, S., Tamang, T.B., D, S.R. and T R, M. (2023) Stock Market Prediction Using Genetic Algorithm Assisted LSTM-CNN Hybrid Model. 2023 *14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, 6-8 July 2023, 1-6. <https://doi.org/10.1109/icccnt56998.2023.10306948>
- [45] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2019) Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding.
- [46] Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., *et al.* (2020) Language Models Are Few-Shot Learners.
- [47] Shen, Y. and Zhang, P.K. (2024) Financial Sentiment Analysis on News and Reports Using Large Language Models and Finbert. 2024 *IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS)*, Shenyang, 26-28 July 2024, 717-721. <https://doi.org/10.1109/icpics62053.2024.10796670>
- [48] Lopez-Lira, A. and Tang, Y. (2023) Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4412788>
- [49] Wang, S.K., Ji, T.R., Wang, L.H., Sun, Y.S., *et al.* (2024) Stock Time: A Time Series Specialized Large Language Model Architecture for Stock Price Prediction.
- [50] Altuner, A.B. and Kilimci, Z.H. (2022) A Novel Deep Reinforcement Learning Based Stock Price Prediction Using Knowledge Graph and Community Aware Sentiments. *Turkish Journal of Electrical Engineering and Computer Sciences*, **30**, 1506-1524. <https://doi.org/10.55730/1300-0632.3862>
- [51] Du, S. and Shen, H. (2024) A Stock Prediction Method Based on Deep Reinforcement Learning and Sentiment Analysis. *Applied Sciences*, **14**, Article 8747. <https://doi.org/10.3390/app14198747>
- [52] Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O. (2017) Proximal Policy Optimization Algorithms.
- [53] Li, Y., Wang, J. and Cao, Y. (2024) A General Framework on Enhancing Portfolio Management with Reinforcement Learning. *Proceedings of the 2024 7th International Conference on Computers in Management and Business*, Singapore, 12-14 January 2024, 119-125. <https://doi.org/10.1145/3647782.3647801>
- [54] Zhang, J. and Lei, Y. (2022) Deep Reinforcement Learning for Stock Prediction. *Scientific Programming*, **2022**, 1-9. <https://doi.org/10.1155/2022/5812546>

- [55] Awad, A.L., Elkaffas, S.M. and Fakhr, M.W. (2023) Stock Market Prediction Using Deep Reinforcement Learning. *Applied System Innovation*, **6**, Article 106. <https://doi.org/10.3390/asi6060106>
- [56] Khan, R.H., Miah, J., Rahman, M.M., Hasan, M.M. and Mamun, M. (2023) A Study of Forecasting Stocks Price by Using Deep Reinforcement Learning. 2023 *IEEE World AI IoT Congress (AIIoT)*, Seattle, 7-10 June 2023, 250-255. <https://doi.org/10.1109/aiiot58121.2023.10174358>
- [57] Deep, A., Shirvani, A., Monico, C., Rachev, S. and Fabozzi, F. (2025) Risk-Adjusted Performance of Random Forest Models in High-Frequency Trading. *Journal of Risk and Financial Management*, **18**, Article 142. <https://doi.org/10.3390/jrfm18030142>