Adaptive and Explainable Margin Trading via Large Language Models on Portfolio Management

Jingyi Gu*
New Jersey Institute of Technology
US
jg95@njit.edu

Junyi Ye*
New Jersey Institute of Technology
US
jy394@njit.edu

Guiling Wang†
New Jersey Institute of Technology
US
gwang@njit.edu

Wenpeng Yin
Penn State University
US
wenpeng@psu.edu

Abstract

Recent strategies for portfolio management often lack flexibility to adjust funds between long and short positions throughout trading periods. This prevents adapting portfolios to the market, which mitigates risks and seizes opportunities. To address these gaps, we propose an adaptive and explainable framework that integrates Large Language Models (LLMs) with Reinforcement Learning (RL) for dynamic long-short position adjustment in response to evolving market conditions. This approach leverages the recent advancements in LLMs for processing unstructured data and their capacity for explainable reasoning. The framework includes two stages: an Explainable Market Forecasting/Reasoning Pipeline, and a Position Reallocation stage. The Market Forecasting/Reasoning Pipeline allows various LLMs to learn market trends from diverse external data sources and determine optimal adjustment ratios with a clear reasoning path. The Portfolio Reallocation stage interacts with the sequential trading process from a pre-trained RL model to enhance decision-making and transparency. Our framework is flexible to accommodate various external data sources from microeconomics to macroeconomics data, diverse data types including time series and news text, along with multiple LLMs. Experiments demonstrate that our framework effectively achieves three times the return and doubles the Sharpe ratio compared to benchmarks. All the data and code are publicly available under NJIT FinTech Lab's GitHub[1].

CCS Concepts

• Applied computing → Economics; **• Computing methodologies → Natural language processing**; **Reinforcement learning**.

*These authors contributed equally to this research.
†Corresponding author.
[1]https://github.com/NJIT-Fintech-Lab/Adaptive-and-Explainable-Margin-Trading-via-LLM-and-RL

ICAIF '24, November 14–17, 2024, Brooklyn, NY, USA

ACM ISBN 979-8-4007-1081-0/24/11
https://doi.org/10.1145/3677052.3698681

Keywords

Portfolio Management, Large Language Model, Market Trend Forecasting, Reinforcement Learning, Explainable AI

ACM Reference Format:
Jingyi Gu, Junyi Ye, Guiling Wang, and Wenpeng Yin. 2024. Adaptive and Explainable Margin Trading via Large Language Models on Portfolio Management. In *5th ACM International Conference on AI in Finance (ICAIF '24), November 14–17, 2024, Brooklyn, NY, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3677052.3698681

1 Introduction

Portfolio management is a crucial aspect of finance, requiring that investors monitor market dynamics and make informed decisions to periodically allocate their portfolios across multiple assets [31]. The goal is to gain profits while maintaining the investment portfolio's stability. Adjusting funds between long and short positions is particularly important as it directly impacts both profit and risk under varying market conditions. Long positions involve buying assets with the expectation of price increases, whereas short positions involve selling borrowed assets to repurchase them at a lower price. In a volatile market, funds in long positions may not be fully invested in assets, therefore reallocating funds from long to short positions can help hedge against potential losses. This dynamic adjustment helps maintain a balanced portfolio, mitigating risks while seizing opportunities during market downturns and upswings. Recent advancements in portfolio management involve strategies that incorporate both long and short positions, often leveraging deep reinforcement learning to adjust the portfolio weights and execute trading operations. However, existing approaches either assume that short positions always maintain equal equity with long positions throughout the trading period, which does not align with real-world conditions, or simply allocate equity to both positions with a fixed long-short equity ratio at initialization, which limits their effectiveness in adapting to rapidly changing environments and achieving better portfolio performance. Therefore, a flexible strategy that incorporates the dynamic adjustment of the long-short position ratio is essential for achieving resilient and adaptive portfolio management, which has been ignored in existing research.

The development of Large Language Models (LLMs) such as GPT-4 [2] and Claude-3 [6] has significantly enhanced data analytics capabilities in complex domains such as finance [38]. LLMs stand out by processing and deriving insights from unstructured data, such as financial reports' textual narratives and social media content, eliminating the need for the extensive preprocessing required by traditional models like BERT. Furthermore, their extensive pre-training enables them to generalize knowledge to new, unseen tasks through zero-shot learning, allowing them to follow natural language instructions and adapt swiftly to the volatile financial markets. LLMs provide contextual explanations for their decisions, boosting transparency and trust among stakeholders, which is crucial in finance to justify investment decisions. They not only augment the analytical capabilities within portfolio management but also contribute to the evolution towards more data-driven and agile financial strategies.

To fill these gaps, we develop an adaptive and explainable framework, integrating LLMs with Reinforcement Learning (RL) to facilitate dynamic fund transfers between two positions in response to evolving market conditions. This framework features a two-stage adjustment process implemented regularly during the RL trading process: *Market Forecasting/Reasoning Pipeline* enables LLMs to interact regularly with RL, leveraging rich external data sources to forecast the fine-grained market trends with clear reasoning and determine the optimal adjustment ratio; In *Position Reallocation stage*, the equity in the state received from RL Environment is adjusted accordingly to reallocate funds between long and short positions, and RL agent continues trading operations with updated portfolio states. We present an example pipeline that utilizes microeconomic firm news and macroeconomic indicator time series, each offering unique insights into general market trends. However, the framework is designed to be highly flexible, capable of incorporating more complex data sources, diverse data types, and various LLMs. Moreover, our designed pipeline not only forecasts the near-future trends and adjustment ratio but also provides a transparent reasoning path with clear explanations, which offers valuable insights into the rationale behind portfolio adjustments and enhances the decision-making process. The framework is conducted on DJIA from 2020/5 to 2024/4 including both bullish and bearish periods. Results demonstrate that intervention of LLM can boost performance at most 3 times of accumulated return (243.852), twice of Sharpe ratio (1.314), and almost threefold Calmar ratio (1.508) compared with backbone without position adjustments, validating the effective adaption capability of LLM to evolving market. Our major contributions are summarized as follows:

(1) We propose the first framework that integrates LLMs interaction with RL for dynamic long-short position adjustments in response to varying market conditions in margin trading.
(2) Our framework offers flexibility to various external data sources from macroeconomics to individual firms, diverse data types including time series and text, as well as multiple LLMs, providing interpretation with varied perspectives.
(3) Our pipeline demonstrates superior performance in enhancing profitability and balancing risks under diverse market conditions.

2 Related Works

2.1 Portfolio Management

A growing body of literature has increasingly explored various strategies and models using reinforcement learning in portfolio management in recent years, driven by the alignment between the sequence decision-making process of trading and the learning process of RL. Early works predominantly focused on long positions. Jiang et al. [14] introduced a Deep Reinforcement Learning framework aiming at maximizing profit, with assumptions on the market fluidity and transactions. Subsequent models like Alpha Stock [23], HRPM [24], and Smart Trader [29] incorporated more realistic trading scenarios, addressing issues such as slippage and interrelationship among assets, and designed specific model structures tailored to these issues. FinRL [16] was provided as an open-access library offering diverse market environments and multiple RL algorithms for various portfolio tasks. However, these works often overlooked short positions for hedging risks in real-world scenarios.

Recent works have begun integrating short positions into trading strategies, such as iRDPG [17], Deep Trader [25], RAT [27], and Meta Trader [18]. However, these approaches either lack detailed settings for short positions, making real-world deployment challenging, or simply assume equal equity between long and short positions along the trading process, without constraints on short positions for risk management. Addressing this gap, Margin Trader [11] emerged as a foundational framework that incorporates realistic constraints to manage the risk and enable leverage trading. While, the initial ratio of long to short positions is set by traders and remains fixed throughout the trading period, motivating us to enhance flexibility by adjusting the ratio to adapt time-varied market conditions.

2.2 LLMs in Finance

The increasing attention on LLMs in the finance sector has led to significant advancements. Unlike traditional machine learning (ML) models trained on specific datasets, LLMs leverage extremely large training datasets, enabling them to perform instruction-following on various tasks with zero-shot capabilities. LLMs have diverse applications in finance, such as automating report/workflow generation [22, 36], predicting market trends[34], analyzing Named Entity Recognition (NER) and sentiment analysis[13, 20, 26, 28, 37], and providing financial advice and question answering[8, 26].

Compared to traditional ML models[10, 12, 32], LLMs offer significant advantages in finance through enhanced reasoning[22], interactivity, and integration capabilities. They provide clear explanations for predictions, improving transparency and trust. Their interactive nature allows dynamic user engagement for refined queries and responses. Moreover, LLMs can seamlessly integrate with various tools [30] and data sources using techniques like Retrieval-Augmented Generation (RAG)[37], delivering comprehensive analyses and decision-making[35]. However, deploying LLMs in finance presents significant challenges. They require immense computational power, leading to environmental concerns. Additionally, they can inherit biases[9] from training data, resulting in unethical outcomes. Their "black-box" nature makes it difficult to understand their conclusions, raising accountability and trust issues. Additionally, small-parameter LLMs often struggle with generalization to new tasks. Currently, they are suited for simpler tasks instead of complex ones like portfolio management. Their integration into advanced financial strategies is limited, needing significant improvements in model capabilities, fine-tuning techniques, and pipeline design. Addressing these issues is crucial for fully harnessing LLMs' potential in finance and ensuring innovative and ethical applications.

3 Preliminary

3.1 Margin Trading

This paper focuses on margin trading with long and short positions. It allows traders to borrow funds from a broker to trade financial assets instead of cash only, enabling them to leverage their positions and potentially amplify their returns. Traders can take long positions if expecting a rising price of an asset, or short positions if betting that the price will fall. In a long position, profits are earned if the asset price increases, whereas, in a short position, profits are realized when the asset's price decreases. As traders profit from positions and equity increases, their buying power including both trader's own funds and borrowed funds is boosted, allowing them to take on larger positions or diversify investments further. Conversely, if trades result in losses, the equity will decrease and the buying power will be reduced. Both long and short positions come with increased risk, as losses are magnified and can exceed the original investment. Therefore, margin trading has constraints to prevent such issues. Brokers require traders to maintain a minimum margin level (typically 40% of the holding market value) and will issue a margin call if the equity in the account falls below this threshold, requiring the trader to deposit additional cash or sell assets to cover the shortfall.

3.2 RL Environment and Long-Short Ratio

Margin trading can be modeled as a Markov Decision Process (MDP) in RL. At the beginning of each time period t, the agent observes the current state s_t. It includes equity conditions on long and short positions separately, as well as the close price, the number of shares held by the agent, and the technical indicators associated with each stock. Then the agent selects an action a_t according to its policy $\pi(a|s)$. The action is defined as a sequence of integer values representing the number of shares to trade for each stock in long and short positions separately. A positive value denotes a buying operation, while a negative value is a selling action. The environment then transitions to the new state s_{t+1} according to transition probability $P(s_{t+1}|s_t, a_t)$, and the agent receives a reward r_t reflecting the immediate benefit of the action. The reward is typically designed to reflect both profits and risks of equity after taking the action from the current step to the next state. Sharpe ratio is typically used as a risk-adjusted return metric in portfolio management, considering both returns and volatility. The objective of RL agent is to learn optimal policy π^* which maximizes expected cumulative reward.

We implement our proposed approach based on an existing RL model, Margin Trader [11], recognized for its state-of-the-art performance in portfolio management. At the initialization step t_0, the agent starts with an initial capital of e. Traders set $r \in [0, 1]$ to determine the proportion of initial equity allocated to long positions relative to the total capital, according to their preference and expectation of the market. Accordingly, $e^l = er$ is allocated to long positions, while $e^s = e(1-r)$ is allocated to short positions. Throughout the trading period, the ratio $r_t = e^l/(e^l + e^s)$ fluctuates as the equity in both long e^l_t and short positions e^s_t adjusts based on market fluctuations and trading operations. However, the allocation of funds between long and short positions remains fixed, meaning capital is not transferred between them despite market changes. It's essential to note that position equity includes both stocks and cash, and funds may not be fully invested. In a bullish market, for example, reallocating more equity from short to long positions rather than holding uninvested cash in shorts can be profitable. Therefore, adapting the long-short ratio and managing uninvested cash based on market conditions and individual stock price movements is crucial for optimizing portfolio performance.

4 Methodology

We aim to develop a flexible framework that dynamically adjusts the long-short ratio in response to time-varied market conditions to enhance the trading strategies, as shown in Figure 1. To achieve this goal, we incorporate state-of-the-art LLMs into the sequential decision-making process of RL. During the trading process, at regular intervals of every k step, the Market Forecasting/Reasoning Pipeline leverages LLMs to analyze complex external data sources to predict near-future asset trends and determine the optimal position ratio change for the upcoming trading periods. Then in the Position Reallocation stage, portfolios are rebalanced between long and short positions to transfer the funds and achieve the updated ratio. The RL agent continues trading operations with the updated portfolio for the next k steps. It is important to note that LLMs intervene only during the deployment phase instead of the training process of RL, thereby our proposed approach advances in efficiency.

4.1 External Datasets

We utilize two distinct datasets to comprehensively analyze the U.S. economic landscape and its interplay with the stock market from different perspectives. The first dataset is a macroeconomic dataset, which includes major indicator time series that reflect broad economic conditions and trends. The second dataset is microeconomic, comprising firm-specific news within the asset pool, providing detailed insights into individual company activities. By examining these datasets, we aim to determine which perspective is more effective for predicting future market changes and enhancing trading strategies, providing a robust analysis of their respective influences on market dynamics.

Macroeconomic Time Series: The dataset comprises monthly time series data of 21 macroeconomic indicators, including CPI (CPI, CPI Less Food and Energy), Interest Rates (Federal Funds Effective Rate, Bank Prime Loan Rate), Government Debt and Deficit (Federal Debt, Federal Surplus or Deficit), Exchange Rates (Nominal Broad US Dollar Index, US Dollar to Euro, Japanese Yen to US Dollar, Chinese Yuan Renminbi to US Dollar), Money Supply (Real M1 Money Stock, Real M2 Money Stock), and Retail Sales (Advance Retail Sales: Retail Trade and Food Services, Advance Retail Sales: Retail Trade). It provides a chronological sequence of data points, enabling the analysis of trends, cycles, and patterns over time, which is crucial for understanding the broader U.S. economic landscape and its potential impact on the stock market.

Microeconomic Firm-Specific News: This dataset is specifically tailored for portfolio management. This dataset focuses on

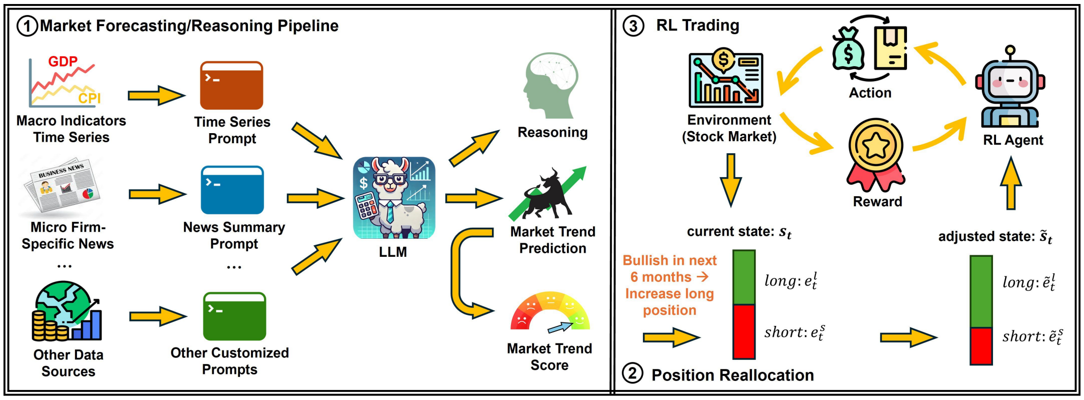

Figure 1: The framework presents the interaction of the two-stage adjustments with the pre-trained RL trading model. Stage 1 (Market Forecasting/Reasoning Pipeline) employs LLMs to predict and analyze future stock market trends, utilizing various data sources. Stage 2 (Position Reallocation) updates the states in the RL process by adjusting the long-short ratio based on the insights gained from LLMs.

daily news articles and reports directly related to individual assets within the portfolio pool, including detailed information on company performance, earnings reports, strategic initiatives, mergers and acquisitions, and other significant corporate events. By capturing firm-specific news, it offers granular insights into the microeconomic factors that influence the performance of major corporations. While macroeconomic news covers general market conditions that may not directly reflect the trends affecting the portfolio pool, firm-specific news provides more relevant and actionable information for managing and optimizing the portfolio.

4.2 Explainable Market Forecasting/Reasoning

We design an explainable market reasoning pipeline via LLM to process the two distinct datasets and generate relevant insights based on corresponding prompts, taking into account the varying data types, data sizes, and the context window limitations of the language models. For each dataset, we have tailored different prompt pipelines. The pipeline leverages the capabilities of advanced language models to analyze, summarize, and interpret financial and market data, providing valuable outputs specific to

Figure 2: Sample prompt during COVID-19 pandemic and GPT-4o response for predicting the overall trend of the DJIA 30 index based on macroeconomic indicator time series.

Table 1: The mapping between market sentiment and expected percentage change over the next 6 months. 6 out of 8 baseline LLMs agreed on this mapping. Phi-3 Medium and DeepSeek-V2 had more conservative ranges, e.g., Strongly Bullish was 10-20% and 15-20% respectively.

Market Sentiment	Expected Percentage Change
Strongly Bullish	Greater than +20%
Bullish	+10% to +20%
Slightly Bullish	+5% to +10%
Flat/Fluctuating	-5% to +5%
Slightly Bearish	-10% to -5%
Bearish	-20% to -10%
Strongly Bearish	Less than -20%

each dataset type. For each data type, LLMs are tasked with predicting the near future (six months) market trend and providing explanations. The market trend prediction is categorized into one of seven fine-grained categories (Strongly Bullish to Strongly Bearish), as illustrated in Table 1. Compared to common categories of positive, negative, and neutral, this allows more nuanced adjustments in subsequent RL algorithms. Below is a detailed description of the LLM pipeline and how it processes each dataset. We use Dow Jones Index Average (DJIA) as a representative example of an asset pool.

4.2.1 Prompt Pipeline for Macroeconomic Indicator Time Series. The prompt in Figure 2 comprises three parts: (1) macroeconomic indicators' definitions, (2) historical macroeconomic indicator time series data, and (3) instructions for LLMs' output requirements. This time series based prompt provides a structured analysis of

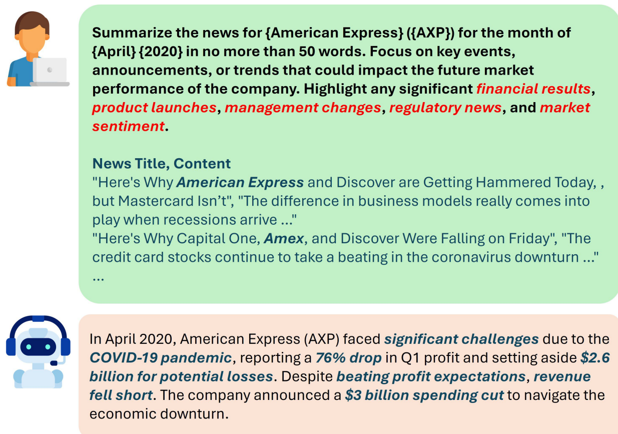

Figure 3: Prompt and GPT-4o response for summarizing the news for American Express (AXP) for the month of April 2020. The texts in the {} are replaced with company names, tickers, month, and year accordingly.

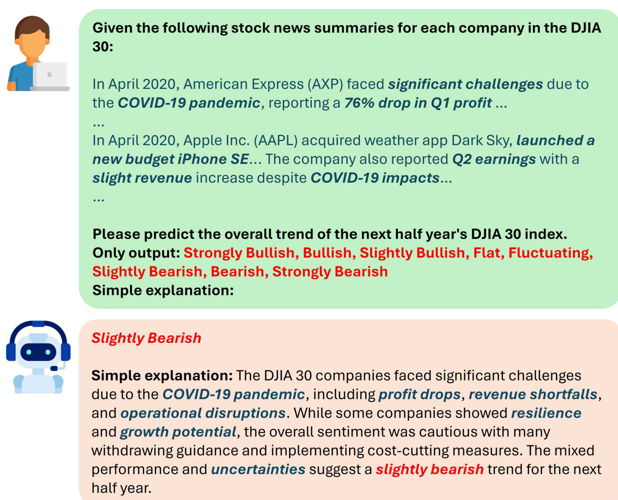

Figure 4: Sample prompt during COVID 19 pandemic and GPT-4o response for predicting the overall trend of the DJIA 30 index based on monthly firm news summary of DJIA 30.

past economic performance by incorporating all relevant indicators. It leverages reliable data from government websites, ensuring comprehensive coverage of economic conditions and trends. By integrating time series data, the pipeline offers a robust framework for understanding the continuous impact of various macroeconomic factors, delivering precise insights into future market trends. It evaluates the LLM's capability to discern relationships between macroeconomic features, detect numerical data trends, and accurately forecast future economic conditions based on the varying sentiments of different indicators.

As shown in Figure 2, GPT-4o analyzes key positive and negative indicators to identify decisive features impacting economic trends. The LLM's reasoning highlights severe economic distress due to the significant drop in GDP, rise in unemployment, decreased inflation and industrial production, increased federal deficit, and plummeting retail sales. Based on this analysis, GPT-4o predicted a strongly bearish market trend during the COVID-19 pandemic, which aligns with the macroeconomic time series data.

4.2.2 Prompt Pipeline for Microeconomic Firm-Specific News. The firm news prompt pipeline involves three key steps to effectively aggregate and summarize daily news for each firm in the pre-selected asset pool, ensuring it fits within the context limits of most open-source LLMs: (1) **Daily Aggregation of Firm News:** For each firm, the daily firm news is collected and aggregated on a monthly basis. This process involves gathering all relevant news articles for each firm, ensuring comprehensive coverage of their activities and developments. (2) **Summarization to Fit LLM Context Limits:** Given the constraint of a 4K-token limit in many LLMs, the aggregated monthly news for each company is summarized to a maximum of 50 words with GPT-4o. This summary focuses on significant financial results, product launches, management changes, regulatory news, and market sentiment, illustrated in Figure 3. This step ensures balanced representation across companies, preventing firms with extensive news coverage (e.g., Apple with over 400 news articles per month) from overshadowing those with fewer updates (e.g., Amgen Inc. with only 30 news articles monthly). By summarizing each company's news, the prompt can accommodate summaries for all 30 companies within the 4K-token window. (3) **Fine-Grained Market Trend Prediction:** The LLM uses the aggregated monthly news summaries for all firms to predict the future fine-grained trend of the market, as illustrated in Figure 4. Shifting the input from macroeconomic data to microeconomic firm news data allows GPT-4o to adjust its forecast from a strongly bearish outlook to a slightly bearish one. This adjustment reflects detailed information in firm news summaries, which indicate resilience and growth potential among companies that have implemented adjustments in response to COVID-19 pandemic. These insights provide more accurate and actionable information not captured in macroeconomic time series.

By focusing on detailed, company-specific information and events, the firm news prompt pipeline provides granular microeconomic insights, offering a more precise perspective for portfolio management. This approach complements the broader economic focus of the macroeconomic pipeline, allowing the LLM to generate nuanced predictions about overall stock market trends based on the significant events of firms in the asset pool.

4.2.3 LLM Response Mapped to Market Trend Score. To seamlessly integrate the predicted trend from LLMs into our RL trading process, qualitative sentiment levels are converted into a quantitative Market Trend Score $m \in \{-3, -2, -1, 0, 1, 2, 3\}$. Positive integers signify varying degrees of bullish sentiment, negative integers represent bearish levels, and zero denotes a flat or fluctuating market condition. A larger absolute value of m indicates a strong market signal. The adjustment ratio $\delta = ms$ is determined by applying a scaling factor s, which controls the impact of trend prediction on position change. For example, if $s = 5\%$, then δ ranges from -15%, to 15%, enabling a precise adjustment based on the market trend prediction provided by LLMs.

4.3 Position Reallocation

The LLM's intervention in the RL process occurs every k step. At time step t, we receive the previous reward r_{t-1} and observe the current state s_t which includes the long equity e_t^l and short equity e_t^s. The current ratio of long equity over the total equity $r_t = e_t^l/(e_t^l + e_t^s)$ can be calculated. We adjust this long ratio proportionally to reach the target ratio $r_t' = r_t * (1 + \delta)$, with the change in long equity being $e_t^l\delta$. Compared with directly adding δ_t on current ratio, this proportional adjustment considers existing equity conditions, thereby preventing significant and abrupt changes in market exposure. In addition, we establish a range threshold for the target ratio $r_t' \in [\eta_1, \eta_2]$, which is set to $[10\%, 90\%]$ in this study, to prevent allocating all funds into a single position. This approach mitigates certain risks of overreacting to market fluctuations and facilitates smoother transitions between positions.

A positive δ results in an increase in long equity, necessitating a transfer of funds from short to long positions. Conversely, a negative δ prompts a movement of funds from long to short. We implement these adjustments following the margin adjustment module in Margin Trader to manage the buying power and states effectively. Firstly, available cash in long positions or available limits for short sales are assessed to cover the transfer amount. If sufficient funds are available, the transfer occurs directly; otherwise, the position closures are triggered, starting with the asset of lowest holding value, as it is considered as the least significant portfolio, until adequate funds are available. In addition, the loan in long or the credit balance in short is also reduced to align with the margin requirement. Secondly, the transferred funds allow for increased short sales or borrowing more loans for the long trade. Consequently, the available limit and credit balance for short positions, or the loan and available cash for long positions, are increased accordingly.

After the adjustment, the state $\tilde{s}_t$ with long equity $\tilde{e}_t^l$ and short equity $\tilde{e}_t^s$ has been updated. The new state is then fed into the RL framework to continue the sequential process and refine the optimal trading policy iteratively.

5 Experiments

5.1 Data Source

Dow Jones Industrial Average (DJIA) is selected as the portfolio pool. We follow [11] for the training, validation, and testing period. Note that the test period is extended to 2020/5 - 2024/2 in our paper, since it is up-to-date and includes complex economic fluctuations, marked by a significant rise (COVID-19 pandemic recovery) and subsequent variations (supply chain disruptions and inflation). The price data of companies in DJIA for RL is sourced from Yahoo Finance[2]. Besides, two distinct external data sources are collected and tested to evaluate their impact on near-future (Six month) US market trend prediction:

(1) **Macroeconomic Indicator Time Series Dataset** comprises monthly time series data for 21 key US economic metrics. The inflation rate data is sourced from the US Inflation Calculator[3], while the remaining macroeconomic indicators are obtained from the Federal Reserve Bank of St. Louis[4]. To ensure consistency, data with daily or quarterly features has been appropriately downsampled or upsampled to a monthly frequency.
(2) **Microeconomic Firm-Specific News Dataset** includes daily news data specific to the 30 companies listed on the DJIA. News is gathered from various sources, such as company announcements, earnings reports, and other significant events. The data is retrieved by ticker from the Stock News API[5], which indexes articles and video content from reputable sources including CNBC, Reuters, MarketWatch, Seeking Alpha, Bloomberg, and The Motley Fool.

5.2 Baseline

To fairly compare the portfolio performance, we employ two benchmarks: **DJIA** follows DJIA index trend, and **Margin Trader** is the backbone RL model without LLM intervention on ratio adjustment. In addition, the following two financial NLP models and several state-of-the-art LLMs are integrated to evaluate our framework: **FinBERT**[7] is an NLP model for financial sentiment analysis, built on BERT and fine-tuned with a comprehensive financial corpus; **FinGPT v3.3**[28] is based on LLaMA2-13B and fine-tuned on financial sentiment datasets; **GPT-4o**[19] excels in multi-language understanding, question answering, math, code evaluation, reasoning, and reading comprehension; **Claude-3.5 Sonnet**[5] significantly outperforms its predecessor, Claude 3 Opus, and competing models across a wide range of evaluations; **DeepSeek-V2**[15] is a highly efficient Mixture-of-Experts model that enables economical training and efficient inference; **Mixtral 8x22B**[3] is a cutting-edge LLM using a Sparse Mixture-of-Experts (SMoE) architecture with reduced computational costs and enhanced speed; **LLaMA-3 70B**[4] is the latest in the LLaMA family, enhancing contextual understanding and efficiency for complex natural language processing tasks; **Qwen-2 72B**[21] is the top model in the Qwen series, handling complex applications effectively; **Yi 34B**[33] excels in English and Chinese benchmarks using advanced architectures like Grouped-Query Attention for versatile applications; **Phi-3 Medium**[1] employs supervised fine-tuning and direct preference optimization for alignment with human preferences and safety.

5.3 Evaluation Metrics

We utilize two primary types of metrics to evaluate portfolio management: profit metrics and risk-adjusted metrics. (1) Profit metric provides a straightforward assessment of portfolio's absolute performance. **Accumulated return** *AR* measuring the overall percentage gain or loss of a portfolio over a specific period. (2) Risk-adjusted metrics consider both returns and risks, providing a more comprehensive evaluation. **Sharpe ratio** *SR* evaluates the excess return per unit of risk. A higher value indicates that portfolio generates more return for each unit of risk taken. **Calmar Ratio** *CR* compares the the return to the largest peak-to-trough loss experienced by the portfolio. A higher value reflects the ability to generate returns while minimizing the significant losses.

[2] https://finance.yahoo.com/
[3] https://www.usinflationcalculator.com/inflation/current-inflation-rates/
[4] https://www.stlouisfed.org/about-us
[5] Stock News API: https://stocknewsapi.com/

Table 2: Portfolio management results. FinBERT and FinGPT, as sentiment analysis models, are unsuitable for macro time series data. FinGPT produces gibberish or multiple outputs with multiple news inputs, making it unsuitable for firm news data.

Model	Scaling factor=10%						Scaling factor=20%					
	Macro Indicator			Firm News			Macro Indicator			Firm News		
	AR(%)	SR	CR	AR(%)	SR	CR	AR(%)	SR	CR	AR(%)	SR	CR
DJI	67.795	0.917	0.646	-	-	-	-	-	-	-	-	-
Margin Trader	75.834	0.673	0.592	-	-	-	-	-	-	-	-	-
FinBERT	-	-	-	55.004	0.558	0.467	-	-	-	-15.561	-0.022	-0.112
FinGPT v3.3	-	-	-	-	-	-	-	-	-	-	-	-
GPT-4o	163.643	1.140	1.077	**205.197**	**1.199**	**1.353**	128.512	1.014	0.657	**243.852**	**1.314**	**1.508**
Claude-3.5 Sonnet	178.530	1.174	1.237	143.189	1.019	0.996	**198.117**	**1.252**	**1.604**	212.229	1.269	1.350
Qwen2 72B	**198.623**	**1.198**	**1.311**	175.395	1.140	1.266	194.996	1.170	1.270	193.232	1.173	1.377
Llama-3 70B	161.010	1.066	1.139	160.060	1.037	1.134	191.449	1.189	1.277	196.969	1.164	1.303
Mixtral 8x22B	160.060	1.037	1.134	160.060	1.037	1.134	196.969	1.164	1.303	196.969	1.164	1.303
DeepSeek-V2	170.553	1.111	1.187	112.173	0.894	0.812	163.889	1.111	1.143	123.893	0.969	0.639
Phi-3 Medium	164.705	1.139	1.192	106.022	0.900	0.920	182.110	1.224	1.494	106.022	0.900	0.920
Yi 34B	160.060	1.037	1.134	82.699	0.724	0.647	196.969	1.164	1.303	82.699	0.724	0.647

5.4 Implementation Details

In our implementation of large language models (LLMs), we employ greedy decoding with a maximum limit of 1024 newly generated tokens. Open-source models are deployed on 1 to 4 A100 GPUs, selected based on GPU RAM requirements, utilizing the Hugging Face library. Proprietary models, such as GPT-4o, Claude 3.5, and DeepSeek V2, are accessed via their respective APIs. To mitigate potential biases in sentiment label assignments, we randomly shuffle the order of sentiment labels in the prompt and conduct five experimental runs. The final prediction is determined through majority voting. For models with a 512-token limit, such as FinBERT and FinGPT, we segment texts into chunks and then apply majority voting. The sentiment outputs—positive, neutral, and negative—are mapped to Bullish, Flat/Fluctuating, and Bearish sentiments, respectively. This approach ensures robust and unbiased sentiment analysis across different model architectures and token limitations.

For the implementation of RL, we choose A2C as the DRL algorithm. For parameters in the RL trading environment and hyperparameters for the A2C, we follow [11]. The initial equity e is $\$100,000$. DJIA uses cash to trade on long positions only. Rest strategies implement both long and short based on Margin Trader, allocating half of the equity at the initialization step. The frequency of interaction (k) between RL and LLMs is a quarter.

5.5 Results and Analysis

5.5.1 Comparison with Baselines. We present two different strategies based on varying scaling factors. Table 2 shows the numerical results for a scaling factor of 10%, representing a common and moderate strategy, and a more radical strategy with a scaling factor of 20%. We have following observations. (1) The frameworks integrating LLMs on a quarterly basis consistently outperform DJIA and Margin Trader without long-short ratio adjustment. This highlights LLMs can effectively detect market trends and determine the position ratio, significantly enhancing portfolio performance. (2) Among LLMs, GPT-4o, Claude-3.5 Sonnet, and Qwen2 72B achieve over twice and even three times the return, a 25% higher Sharpe ratio, and more than twice the Calmar ratio across datasets and scaling factors compared with DJI and Margin Trader, making them standout LLMs for this framework. This conclusion aligns with the performance of these LLMs in other tasks. In contrast, FinBERT exhibits suboptimal performance, and FinGPT often generates gibberish or multiple outputs, rendering it unsuitable for this application. This issue arises because both models are trained on individual news inputs, leading to disorganized outputs. This limitation highlights a fundamental drawback of fine-tuned NLP models on isolated tasks: their lack of generalization. Unlike LLMs, which are designed to handle diverse and complex data inputs, fine-tuned models struggle to generalize beyond their specific training data, resulting in less reliable performance in varied contexts. (3) Increasing the scaling factor from 10% to 20% results in a more significant ratio change and generally enhances the performance of most models. However, it is important to note that larger scaling factors also introduce higher risk and greater volatility in equity changes, such as Llama-3 70B in firm news and Deepseek-V2 in macro indicators. While this may boost performance in a bullish market, it can also lead to extreme position changes and excessive market exposure, increasing the risk of substantial losses in unexpected market conditions. (4) Generally, models perform better with macro indicators than firm news, as macro indicators offer extensive information on the broader economic environment. Prices of individual assets already reflect some impact of their firm-specific news, potentially reducing its additional predictive power. An exception is GPT-4o, which excels at capturing and interpreting the nuanced and immediate signals in firm news. Note that our framework offers flexibility for traders to choose data sources for ratio adjustments, ensuring informed decision-making even when certain data sources are unavailable.

5.5.2 Effectiveness of LLM Predictions on Market Trends. To illustrate the effectiveness of predictions of LLM on market trends, Figure 5 compares trend predictions by GPT-4o with DJIA movements from the test period. The DJIA close prices are shown in blue, while sentiment predictions are indicated by colored blocks: more red denotes bearish sentiment, and more green denotes bullish sentiment.

The plot demonstrates the alignment of GPT-4o's predictions with actual market trends.

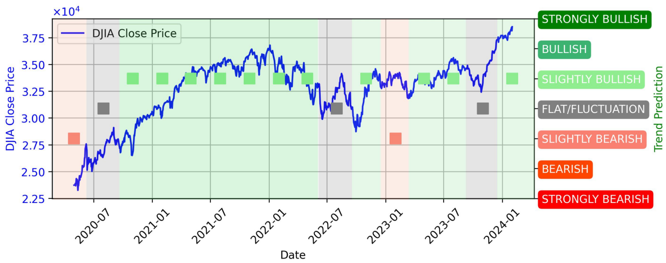

Figure 5: Comparison of GPT-4o trend predictions using micro firm-specific news as input with DJIA movements.

5.5.3 Frequency Analysis. To investigate the impact of frequency of adjusting the long-short ratio based on LLM output, we implement three different frequencies: quarterly, bi-annually, and annually. Figure 6 presents the equity changes of Claude in varying frequencies. Firstly, quarterly adjustments introduce higher volatility and lower returns compared to less frequent adjustments such as bi-annual ones, because macro information and firm news tend to have long-term effects, and adjusting positions too frequently makes portfolios unstable and unbalanced. Besides, infrequent adjustments, such as annual adjustments, show lower performance, missing out on opportunities to capitalize on timely market movements and thus providing suboptimal portfolio performance. Therefore, it is important to find a reasonable frequency for optimal portfolio performance. Bi-annual adjustments strike the optimal balance, avoiding the excessive volatility of quarterly adjustments and the missed opportunities of annual adjustments, achieving the highest cumulative return, sharpe ratio, and calmar ratio, particularly excelling with macro indicators and firm news.

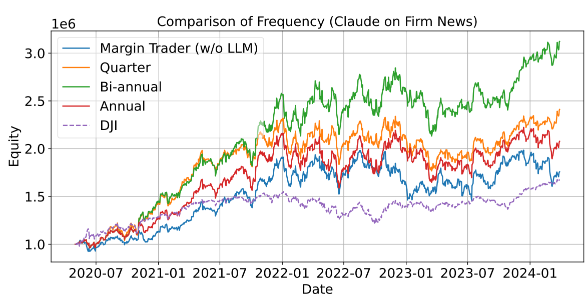

Figure 6: Equity changes on various adjustment frequencies.

6 Conclusion

Our study presents a novel framework that dynamically adjusts the long-short position ratio in portfolio management by leveraging the interaction between LLMs and RL. This framework ensures timely adaptation to evolving market conditions, offering transparency and informed decision-making through its explainable market reasoning module. Additionally, it provides traders with flexibility by enabling the integration of diverse external data sources and various LLMs. Comprehensive evaluations on real-world markets demonstrate that the intervention of LLMs significantly improves trading strategies in terms of both profitability and risk management. These findings underscore the transformative potential of LLMs and RL in revolutionizing portfolio management, paving the way for more adaptive, robust, and effective financial strategies.

Acknowledgments

We would like to express our sincere gratitude to Travis Virgil, Neo Maralit, and Jean Arly Volmar for their contributions to the data collection process in support of this paper.

References

[1] Marah Abdin, Sam Ade Jacobs, Ammar Ahmad Awan, Jyoti Aneja, Ahmed Awadallah, Hany Awadalla, Nguyen Bach, Amit Bahree, Arash Bakhtiari, Harkirat Behl, et al. 2024. Phi-3 technical report: A highly capable language model locally on your phone. *arXiv preprint arXiv:2404.14219* (2024).

[2] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

[3] Mistral AI. 2024. Cheaper, Better, Faster, Stronger. https://mistral.ai/news/mixtral-8x22b/. Accessed: 2024-07-15.

[4] Meta AI. 2024. Introducing Meta Llama 3: The Most Capable Openly Available LLM to Date. https://ai.meta.com/blog/meta-llama-3/. Accessed: 2024-07-15.

[5] Anthropic. 2024. Introducing Claude 3.5 Sonnet. https://www.anthropic.com/news/claude-3-5-sonnet. Accessed: 2024-07-15.

[6] AI Anthropic. 2024. The claude 3 model family: Opus, sonnet, haiku. *Claude-3 Model Card* 1 (2024).

[7] Dogu Araci. 2019. Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063* (2019).

[8] Xiang Deng, Vasilisa Bashlovkina, Feng Han, Simon Baumgartner, and Michael Bendersky. 2023. What do llms know about financial markets? a case study on reddit market sentiment analysis. In *Companion Proceedings of the ACM Web Conference 2023*. 107–110.

[9] Isabel O Gallegos, Ryan A Rossi, Joe Barrow, Md Mehrab Tanjim, Sungchul Kim, Franck Dernoncourt, Tong Yu, Ruiyi Zhang, and Nesreen K Ahmed. 2024. Bias and fairness in large language models: A survey. *Computational Linguistics* (2024).

[10] Jingyi Gu, Fadi P Deek, and Guiling Wang. 2023. Stock broad-index trend patterns learning via domain knowledge informed generative network. *arXiv preprint arXiv:2302.14164* (2023).

[11] Jingyi Gu, Wenlu Du, AM Muntasir Rahman, and Guiling Wang. 2023. Margin trader: a reinforcement learning framework for portfolio management with margin and constraints. In *Proceedings of the Fourth ACM International Conference on AI in Finance*. 610–618.

[12] Jingyi Gu, Wenlu Du, and Guiling Wang. 2024. RAGIC: Risk-Aware Generative Adversarial Model for Stock Interval Construction. *arXiv preprint arXiv:2402.10760* (2024).

[13] Jingyi Gu, Sarvesh Shukla, Junyi Ye, Ajim Uddin, and Guiling Wang. 2023. Deep learning model with sentiment score and weekend effect in stock price prediction. *SN Business & Economics* 3, 7 (2023), 119.

[14] Zhengyao Jiang, Dixing Xu, and Jinjun Liang. 2017. A deep reinforcement learning framework for the financial portfolio management problem. *arXiv preprint arXiv:1706.10059* (2017).

[15] Wenfeng Liang, Aixin Liu, Bei Feng, et al. 2024. DeepSeek-V2: A Strong, Economical, and Efficient Mixture-of-Experts Language Model. arXiv:2405.04434 [cs.CL]

[16] Xiao-Yang Liu, Hongyang Yang, Qian Chen, Runjia Zhang, Liuqing Yang, Bowen Xiao, and Christina Dan Wang. 2020. FinRL: A deep reinforcement learning library for automated stock trading in quantitative finance. *arXiv preprint arXiv:2011.09607* (2020).

[17] Yang Liu, Qi Liu, Hongke Zhao, Zhen Pan, and Chuanren Liu. 2020. Adaptive quantitative trading: An imitative deep reinforcement learning approach. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 34. 2128–2135.

[18] Hui Niu, Siyuan Li, and Jian Li. 2022. MetaTrader: An reinforcement learning approach integrating diverse policies for portfolio optimization. In *Proceedings of the 31st ACM international conference on information & knowledge management*.

[19] OpenAI. 2024. Hello GPT-4o. https://openai.com/index/hello-gpt-4o/. Accessed: 2024-07-15.

[20] AM Muntasir Rahman, Ajim Uddin, and Guiling Grace Wang. 2023. HODL: The Hold of Reddit Over the Stock Market. In *2023 IEEE 39th International Conference*

on Data Engineering Workshops (ICDEW). IEEE, 36–39.

[21] Qwen Team. 2024. Hello Qwen2. https://qwenlm.github.io/blog/qwen2/. Accessed: 2024-07-15.

[22] Ana Clara Teixeira, Vaishali Marar, Hamed Yazdanpanah, Aline Pezente, and Mohammad Ghassemi. 2023. Enhancing Credit Risk Reports Generation using LLMs: An Integration of Bayesian Networks and Labeled Guide Prompting. In *Proceedings of the Fourth ACM International Conference on AI in Finance*. 340–348.

[23] Jingyuan Wang, Yang Zhang, Ke Tang, Junjie Wu, and Zhang Xiong. 2019. Alphastock: A buying-winners-and-selling-losers investment strategy using interpretable deep reinforcement attention networks. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*.

[24] Rundong Wang, Hongxin Wei, Bo An, Zhouyan Feng, and Jun Yao. 2021. Commission fee is not enough: A hierarchical reinforced framework for portfolio management. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

[25] Zhicheng Wang, Biwei Huang, Shikui Tu, Kun Zhang, and Lei Xu. 2021. DeepTrader: a deep reinforcement learning approach for risk-return balanced portfolio management with market conditions Embedding. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 35. 643–650.

[26] Shijie Wu, Ozan Irsoy, Steven Lu, Vadim Dabravolski, Mark Dredze, Sebastian Gehrmann, Prabhanjan Kambadur, David Rosenberg, and Gideon Mann. 2023. Bloomberggpt: A large language model for finance. *arXiv preprint arXiv:2303.17564* (2023).

[27] Ke Xu, Yifan Zhang, Deheng Ye, Peilin Zhao, and Mingkui Tan. 2021. Relation-aware transformer for portfolio policy learning. In *Proceedings of the twenty-ninth international conference on international joint conferences on artificial intelligence*.

[28] Hongyang Yang, Xiao-Yang Liu, and Christina Dan Wang. 2023. FinGPT: Open-Source Financial Large Language Models. *FinLLM Symposium at IJCAI 2023* (2023).

[29] Mengyuan Yang, Xiaolin Zheng, Qianqiao Liang, Bing Han, and Mengying Zhu. 2022. A Smart Trader for Portfolio Management based on Normalizing Flows.. In *IJCAI*. 4014–4021.

[30] Junyi Ye, Mengnan Du, and Guiling Wang. 2024. DataFrame QA: A Universal LLM Framework on DataFrame Question Answering Without Data Exposure. *arXiv preprint arXiv:2401.15463* (2024).

[31] Junyi Ye, Bhaskar Goswami, Jingyi Gu, Ajim Uddin, and Guiling Wang. 2024. From Factor Models to Deep Learning: Machine Learning in Reshaping Empirical Asset Pricing. *arXiv preprint arXiv:2403.06779* (2024).

[32] Junyi Ye, Jingyi Gu, Ankan Dash, Fadi P Deek, and Guiling Grace Wang. 2023. Prediction with time-series mixer for the S&P500 index. In *2023 IEEE 39th International Conference on Data Engineering Workshops (ICDEW)*. IEEE, 20–27.

[33] Alex Young, Bei Chen, Chao Li, Chengen Huang, Ge Zhang, Guanwei Zhang, Heng Li, Jiangcheng Zhu, Jianqun Chen, Jing Chang, et al. 2024. Yi: Open foundation models by 01. ai. *arXiv preprint arXiv:2403.04652* (2024).

[34] Xinli Yu, Zheng Chen, and Yanbin Lu. 2023. Harnessing LLMs for Temporal Data: A Study on Explainable Financial Time Series Forecasting. In *Proceedings of the EMNLP 2023 Industry Track*. Association for Computational Linguistics.

[35] Yangyang Yu, Haohang Li, Zhi Chen, Yuechen Jiang, Yang Li, Denghui Zhang, Rong Liu, Jordan W Suchow, and Khaldoun Khashanah. 2024. FinMem: A performance-enhanced LLM trading agent with layered memory and character design. In *Proceedings of the AAAI Symposium Series*, Vol. 3. 595–597.

[36] Zhen Zeng, William Watson, Nicole Cho, Saba Rahimi, Shayleen Reynolds, Tucker Balch, and Manuela Veloso. 2023. FlowMind: automatic workflow generation with LLMs. In *Proceedings of the Fourth ACM International Conference on AI in Finance*. 73–81.

[37] Boyu Zhang, Hongyang Yang, Tianyu Zhou, Muhammad Ali Babar, and Xiao-Yang Liu. 2023. Enhancing financial sentiment analysis via retrieval augmented large language models. In *Proceedings of the fourth ACM international conference on AI in finance*. 349–356.

[38] Huaqin Zhao, Zhengliang Liu, Zihao Wu, Yiwei Li, Tianze Yang, Peng Shu, Shaochen Xu, Haixing Dai, Lin Zhao, Gengchen Mai, et al. 2024. Revolutionizing finance with llms: An overview of applications and insights. *arXiv preprint arXiv:2401.11641* (2024).