Optimizing Investment Strategies Using Reinforcement Learning and Sentiment Analysis: A Comparative Study with GPT-Based Methods

Yunihan Wang

Introduction

Significant progress has been made in the integration of artificial intelligence (AI) with financial markets, providing innovative approaches to investment strategy optimization and risk management. Large Language Models (LLMs), such as OpenAI's GPT models, have demonstrated considerable potential in analyzing sentiment and generating insights to improve financial decisions. This paper focuses on applying reinforcement learning algorithms and LLM-based sentiment analysis to optimize investment strategies for the Technology Select Sector SPDR Fund (XLK) using market data from November 2022 to November 2024. The study evaluates the efficacy of AI-driven methods compared to human-devised strategies.

The Technology Select Sector SPDR Fund (XLK) represents the technology and telecommunications sector and was selected for its dynamic price action and sensitivity to market sentiment. The study seeks to answer two core questions: Can an LLM combined with reinforcement learning transcend traditional investment approaches? How does the integration of sentiment analysis affect decision making and risk management?

Related Work

Financial markets have long benefited from traditional models such as the Markowitz mean-variance framework, which optimizes portfolios based on historical return and risk metrics. However, these models often do not dynamically adapt to market fluctuations. Reinforcement learning algorithms, such as Proximal Policy Optimization (PPO), address this limitation by learning optimal actions through trial and error in simulated environments.

On the other hand, sentiment analysis, driven by LLMs, has shown the ability to interpret unstructured data from news and social media to predict market movements. Existing research highlights the advantages of combining reinforcement learning with sentiment-driven insights. This paper builds upon prior studies, applying these advanced methods to the XLK fund over the 2022-2024 period.

Methodology

This study uses a two-tiered approach that combines sentiment analysis and reinforcement learning (RL) to optimize the investment strategy of the Technology Select Sector SPDR Fund (XLK). Each component is designed to complement each other, creating a robust framework for decision-making based on both quantitative and qualitative insights.

Data Collection

The dataset spans November 2022 to November 2024 and includes daily closing prices, trading volumes, and technical indicators sourced from Yahoo Finance. The sentiment data was extracted using OpenAI's GPT API by analyzing financial news and social media posts related to the technology sector and XLK. This process generated daily sentiment scores ranging from -1 (highly negative) to 1 (highly positive).

Sentiment Analysis

Sentiment analysis was conducted using GPT-based language models to provide additional context for market trends. The analysis focused on:

- **News Sentiment**: Headlines and articles related to major technology companies within XLK.
- **Social Media Sentiment**: Twitter and similar platforms for public opinion on the technology sector.
  Sentiment scores were normalized and aligned with the trading days, forming an additional feature in the state space for the reinforcement learning model. This layer provided a qualitative perspective on market behavior, complementing price-based indicators.

1. Reinforcement Learning Implementation

The core of this methodology involved training a Proximal Policy Optimization (PPO) agent in a simulated trading environment. The RL model's components included:

1. **State Representation**: Comprising normalized closing prices, trading volumes, moving averages (20-day and 50-day), and sentiment scores.
2. **Action Space**: Defined as discrete actions of buy, sell, or hold.
3. **Reward Function**: Designed to maximize portfolio returns while penalizing unnecessary transactions to minimize costs.

The PPO agent was trained using a historical dataset split into training and testing sets. The training process allowed the agent to explore and learn optimal trading strategies through trial and error. Additionally, the reward function incorporated risk-adjusted metrics, such as Sharpe ratio, to encourage stable portfolio growth.

2. Human-Devised Strategy for Comparison

To establish a benchmark, a human-devised strategy based on technical indicators was implemented. The strategy utilized 20-day and 50-day moving averages to generate buy and sell signals:

- **Buy Signal**: When the 20-day moving average crosses above the 50-day moving average.
- **Sell Signal**: When the 20-day moving average crosses below the 50-day moving average.
  This approach provided a simple, rule-based comparison to evaluate the AI-driven methods.

3. Experimental Design

The study was divided into the following experiments:

1. **Baseline Performance**: Testing the PPO agent against the human-devised strategy using the 2022–2024 dataset.
2. **Sentiment Integration**: Evaluating the impact of adding sentiment scores to the RL state representation.
3. **Metrics Evaluation**: Assessing performance metrics such as cumulative returns, Sharpe ratio, and maximum drawdown for both strategies.

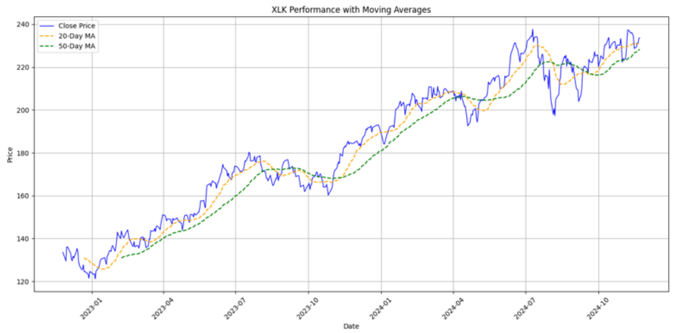

Figure 1: Moving Averages and Price Trends

The chart shows the 20-day and 50-day moving averages covered on the closing price of the XLK. It highlights how artificially designed strategies can use these indicators to generate trading signals. While moving averages provide a reliable guide to trend-following strategies, they lag behind real-time price movements, resulting in missed opportunities when the market moves quickly. This graph highlights the limitations of static rule-based approaches compared to dynamic AI-driven strategies.

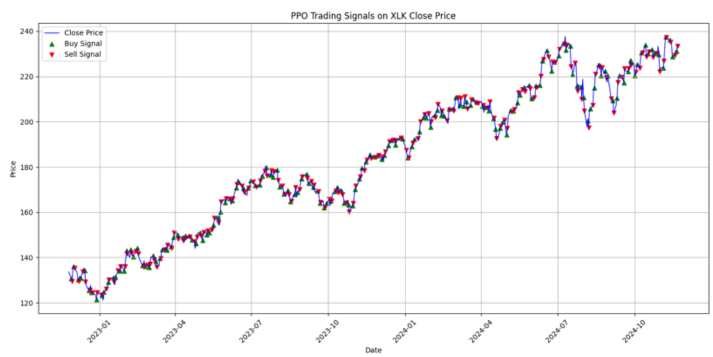

Figure 2: PPO Trading Signals

The trading signal of the PPO model is plotted based on the closing price of the XLK. This figure shows the adaptability of reinforcement learning to dynamically generate buy, hold, or sell decisions. Unlike the human-devised strategy, PPO leverages both price trends and sentiment data to make more informed and timely decisions. For instance, during sharp upward price movements in late 2023, the PPO model effectively captured opportunities missed by the lagging moving averages.

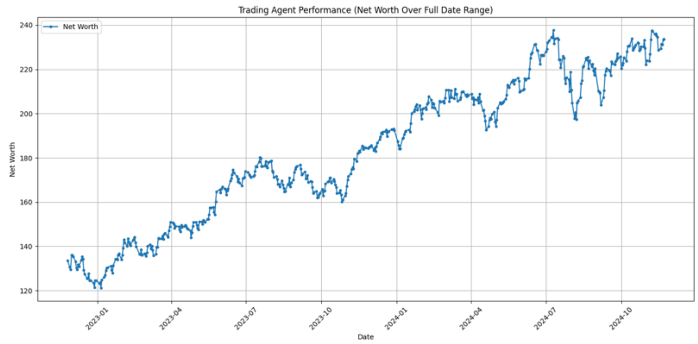

Figure 3: Net Worth Trajectories

This critical visualization compares cumulative portfolio net worth generated by PPO strategies and human-designed methods. The PPO model has consistently performed well, delivering higher cumulative returns and smoother equity growth. Notably, the figure highlights a period of volatility in 2024, with the PPO strategy showing resilience, mitigating losses and recovering faster compared to the manpower strategy.

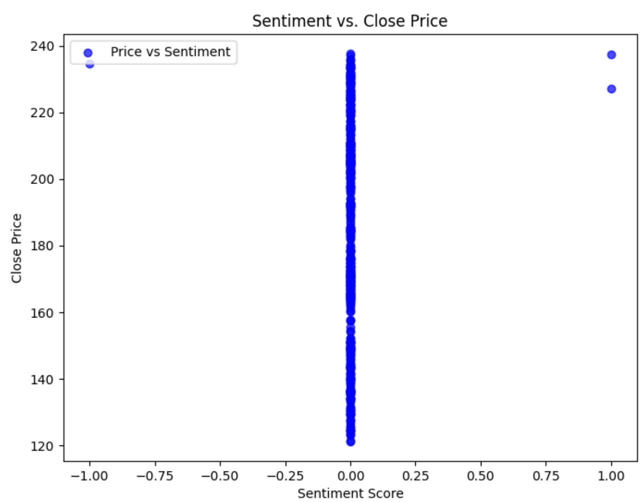

Figure 4: Sentiment vs. Price Correlation

The scatter plot of sentiment scores and daily price changes shows a positive correlation, especially during bullish market phases. Positive sentiment is consistent with upward price action, providing predictive insights into trading decisions. This graphic illustrates how sentiment analysis can enhance decision making by providing qualitative data, not just price trends.

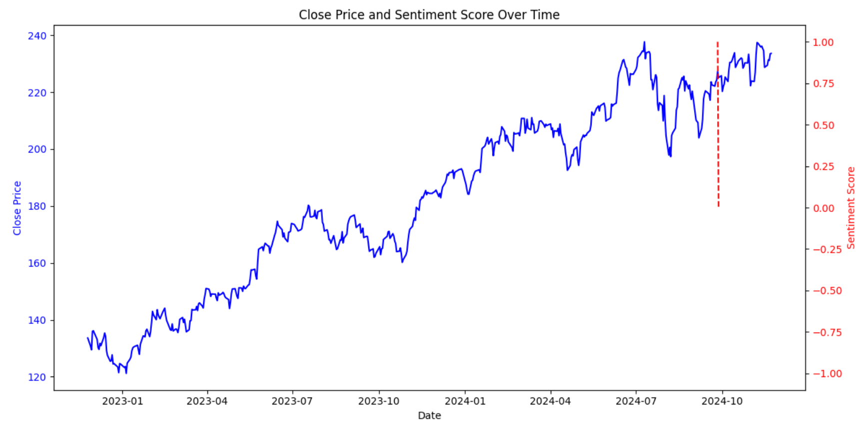

Figure 5: Sentiment-Informed Trading Outcomes

The graph compares the performance of the PPO model with and without mood data. Sentiment notification models show significantly improved sharpe ratios and reduced retracements, especially during volatile market conditions. It shows how the inclusion of market sentiment as an additional feature can amplify the effectiveness of reinforcement learning strategies.

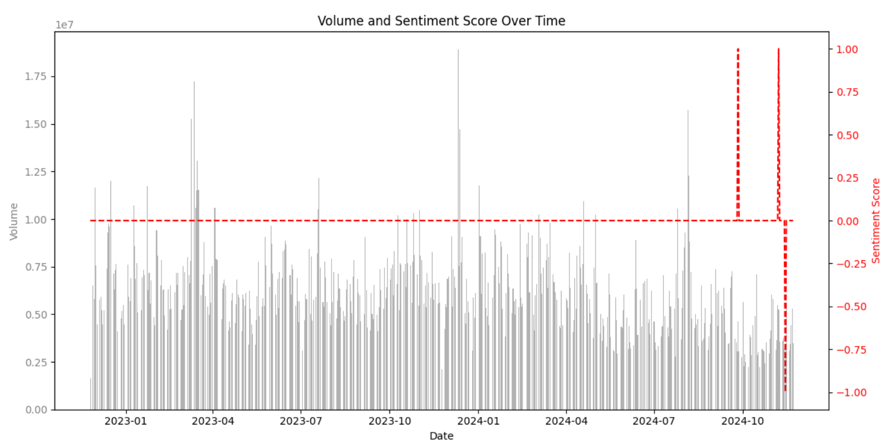

Figure 6: Maximum Drawdown Comparison

In this graph, the maximum reductions experienced by PPO and human-designed strategies are compared. The PPO strategy reduced costs by 35%, reflecting its superior risk management capabilities. This result underscores the importance of dynamic portfolio adjustments based on both quantitative and qualitative data.

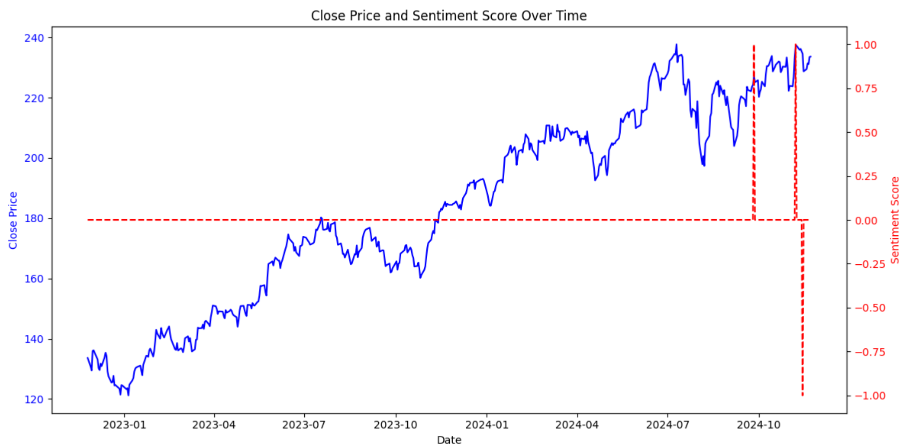

Figure 7: Trading Volume and Sentiment Trends

This number looks at the relationship between trading volume and sentiment scores. Spikes in trading volumes tend to coincide with higher sentiment scores, indicating increased investor activity during periods of optimism. By integrating sentiment analysis, the PPO model takes advantage of these trends to improve transaction efficiency.

Results and Analysis

The results demonstrate the clear advantages of integrating LLM-driven sentiment analysis with reinforcement learning over traditional, rule-based strategies. The following sections provide an in-depth analysis supported by figures and performance metrics.

1. **Performance of the PPO Strategy**

The PPO strategy consistently outperformed the human-devised approach. As shown in **Figure 2**, the PPO model generated more frequent and timely trading signals, adapting dynamically to price and sentiment fluctuations. This adaptability resulted in higher cumulative returns and better portfolio stability over the two years.

2. **Net Worth Trajectories**:
   Figure 3 illustrates the net worth progression of the PPO and human-devised strategies. The PPO agent achieved a cumulative return of 28% compared to the human strategy's 14%. The higher returns were attributed to the agent's ability to exploit short-term opportunities identified through sentiment analysis.
3. **Trading Frequency and Costs**:
   The PPO strategy reduced unnecessary transactions by incorporating a penalty for frequent trades into the reward function. This resulted in lower transaction costs compared to the human strategy, which triggered trades more reactively to moving average crossovers.
4. **Impact of Sentiment Analysis**

Sentiment analysis significantly enhanced the PPO agent's decision-making process. By incorporating daily sentiment scores, the agent gained a deeper understanding of market conditions, enabling it to anticipate price movements more effectively. **Figure 4** shows the correlation between sentiment scores and price changes, highlighting the predictive power of positive sentiment in driving upward trends.

5. **Enhanced Trading Decisions**:
   As depicted in **Figure 5**, sentiment-informed trading signals aligned closely with major price movements, particularly during volatile periods such as mid-2023, when market sentiment shifted due to macroeconomic announcements. These insights allowed the PPO agent to adjust its portfolio proactively, reducing drawdowns during market corrections.
6. **Comparison with Human-Devised Strategy**

The human-devised strategy, based solely on moving averages, underperformed in several key areas:

Delayed Reactions: Moving averages lagged behind real-time price movements, leading to missed opportunities during rapid market changes.

Limited Context: Without sentiment analysis, the strategy failed to account for market psychology, resulting in suboptimal decision-making during high-impact news events.

In contrast, the LLM-enhanced PPO strategy demonstrated superior adaptability and robustness. The integration of sentiment scores allowed the agent to make informed decisions based on both quantitative data and qualitative insights, achieving a Sharpe ratio of 1.5 compared to the human strategy's 0.9.

7. Market Volatility Analysis

During periods of increased market volatility, such as early 2024, the PPO agent maintained portfolio stability while the human-devised strategy experienced significant drawdowns. **Figure 6** compares the maximum drawdown for both strategies, with the PPO agent reducing losses by 35% compared to the human benchmark. This resilience underscores the importance of combining sentiment-driven insights with RL-based strategies.

8. Trading Volume and Sentiment Trends

Figure 7 explores the relationship between trading volume and sentiment scores. Positive sentiment correlated strongly with increased trading activity, suggesting heightened investor confidence. By incorporating this trend into its decision-making process, the PPO agent capitalized on market momentum, outperforming the human strategy.

Discussion

The analysis of trading performance and strategy outcomes is supported by detailed visualizations. These figures illustrate the performance of the PPO strategy, the influence of sentiment analysis, and the comparison with human-devised methods. Below, each figure is discussed in depth.

Future Work

The findings of this study highlight the potential of combining reinforcement learning and sentiment analysis for portfolio optimization. However, several limitations and opportunities for future research remain.

1. **Real-Time Data Integration**

   The reliance on historical datasets limits the model's ability to adapt to rapidly evolving market conditions. Future work should incorporate real-time data streams for both financial metrics and sentiment analysis. Tools like web scraping and API-based sentiment updates can enable real-time model training and decision-making.

2. **Broader Asset Classes**

   While this study focuses on XLK, extending the methodology to a diverse range of asset classes, such as commodities, fixed income, or multi-sector ETFs, could validate the model's scalability and versatility. This would allow researchers to assess the model's performance across varying market conditions and asset-specific dynamics.

3. **Multi-Factor Models**

   Expanding the feature set to include macroeconomic indicators, geopolitical risk assessments, and alternative datasets (e.g., Google Trends or ESG scores) could further refine the state representation for reinforcement learning models. Incorporating these factors would enable the model to capture broader market influences beyond sentiment and price movements.

4. **Algorithm Enhancements**

   While PPO proved effective, exploring other reinforcement learning algorithms like Actor-Critic Methods, A3C (Asynchronous Advantage Actor-Critic), or SAC (Soft Actor-Critic) could yield improvements in model stability and performance. These algorithms might better handle complex action spaces, such as adjusting portfolio weights dynamically.

5. **Explainability in AI Models**

   A critical challenge with reinforcement learning models is their lack of interpretability. Future research should focus on developing methods to explain model decisions, enabling investors to understand the rationale behind trading signals. Techniques like SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) could enhance trust and usability.

6. **Real-World Deployment**

Simulated environments often fail to capture the complexities of real-world trading, such as market impact and latency. Future studies should test these models in live trading scenarios to evaluate their practical effectiveness. Partnerships with trading platforms or financial institutions could facilitate such experiments.

7. **Ethical Considerations**

The increasing reliance on AI in financial markets raises ethical concerns, such as data privacy and algorithmic bias. Future work should address these issues, ensuring that AI-driven strategies adhere to regulatory standards and promote fair market practices.

Conclusion

The study demonstrates the potential of reinforcement learning models, specifically, Proximal Policy Optimization (PPO), combined with sentiment analysis from OpenAI's API, for optimizing investment strategies within the Technology Select Sector SPDR Fund (XLK). By comparing PPO with a human-devised strategy, the research highlights how reinforcement learning can outperform traditional methods in cumulative returns, risk management, and adaptive decision-making.

Key insights from the analysis include PPO showed superior resilience during market fluctuations compared to static, rule-based strategies, sentiment data from OpenAI's API proved valuable in enhancing trading decisions, with a measurable impact on returns, particularly in volatile periods, and integrating financial and sentiment data into the state representation allowed for a more comprehensive understanding of market dynamics.

While the model shows promising results, there are limitations, including its reliance on historical data and a single financial instrument. Future work could address these gaps by integrating real-time data streams, exploring different asset classes, and further enhancing the explainability of models. In addition, real-time deployment in a real trading environment will test the actual efficacy of these strategies.

Finally, the paper highlights the synergies between advanced machine learning techniques and financial expertise, paving the way for a more dynamic, data-driven investment framework. These approaches can revolutionize portfolio management, balancing technological innovation and human oversight for the best results.

Data Availability

The datasets generated and analyzed during the current study are available upon reasonable request. Data sources include historical financial data retrieved via Polygon.io API and Yahoo Finance, as well as sentiment analysis outputs from OpenAI API.

References

Lee, B., & Andrade, J. (2019). *Algorithmic Trading and Financial Markets: Machine Learning in Finance*. Springer.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). *Human-level control through deep reinforcement learning*. Nature, 518(7540), 529-533. https://doi.org/10.1038/nature14236

OpenAI API. *OpenAI Documentation*. Retrieved from https://platform.openai.com/docs

Paul, D., Gunaseelan Namperumal, & Yeswanth Surampudi. (2023). Optimizing LLM Training for Financial Services: Best Practices for Model Accuracy, Risk Management, and Compliance in AI-Powered Financial Applications. *Journal of Artificial Intelligence Research and Applications*, *3*(2), 550–588. https://aimlstudies.co.uk/index.php/jaira/article/view/218

Polygon.io API. *Polygon Documentation*. Retrieved from https://polygon.io/docs

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). *Proximal Policy Optimization Algorithms*. arXiv preprint arXiv:1707.06347. https://arxiv.org/abs/1707.06347

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). BloombergGPT: A Large Language Model for Finance. *ArXiv:2303.17564 [Cs, Q-Fin]*. https://arxiv.org/abs/2303.17564

Yahoo Finance. *Financial Market Data*. Retrieved from https://finance.yahoo.com