**Sentiment-Driven Cryptocurrency Trading Using FinBERT and FINMEMAgent**

**Author:** Team 6 –

SABARI MUKUNDTH JAYARAM,

OMKAR JADHAV,

ANUSHAKA CHOUGULE,

JAY SONAGRA

**Course Information**

**Professor's Name:** Dr. Liao

**Course:** AIT 526 - Data Mining - Section 001

**University:** George Mason University

**Abstract**

This project investigates the integration of sentiment analysis on financial news to inform cryptocurrency trading strategies. Leveraging FinBERT, a pre-trained financial sentiment analysis model, we fine-tuned it using the Financial Phrase Bank dataset through a cross-validation approach to enhance its performance on financial texts. We developed FINMEMAgent, a trading agent that processes news headlines related to Bitcoin (BTC) and Ethereum (ETH), analyzes sentiments, and makes trading decisions (BUY, SELL, HOLD) based on sentiment scores and technical indicators such as the Simple Moving Average (SMA). A comprehensive trading simulation was conducted using historical data, evaluating the agent's performance in terms of portfolio value, Sharpe ratio, and maximum drawdown. The findings suggest that sentiment analysis can provide valuable insights for trading strategies, although challenges like overfitting during model fine-tuning and data sparsity in news headlines persist. This report details the methodologies employed, results obtained, challenges encountered, and potential avenues for future enhancements.

**Keywords**

* Cryptocurrency Trading
* Sentiment Analysis
* FinBERT
* Machine Learning
* Financial News
* Natural Language Processing (NLP)
* FINMEMAgent
* Trading Strategy

**Introduction**

The explosive growth of cryptocurrencies has captured the attention of both investors and researchers, presenting unique opportunities and challenges due to their inherent volatility and the continuous nature of their trading markets. Traditional trading strategies predominantly rely on technical analysis of historical price data. However, the advent of information-driven markets underscores the significant impact that news and social media sentiments have on market dynamics. Sentiment analysis of financial news offers a mechanism to quantify market sentiment, potentially enhancing trading strategies by providing real-time insights.

This project aims to synergize sentiment analysis of financial news with conventional technical indicators to develop an adaptive trading agent, FINMEMAgent. By utilizing FinBERT—a transformer-based model pre-trained on financial text—we analyze news headlines related to cryptocurrencies and execute informed trading decisions. The primary objective is to evaluate whether the incorporation of sentiment analysis can augment trading performance within the volatile cryptocurrency market.

**Motivation & Background**

Cryptocurrencies are renowned for their price volatility, which is often influenced by market sentiment in response to news events. Traditional trading strategies that rely solely on technical indicators may not adequately capture the immediate effects of news on price movements. Sentiment analysis provides a quantitative measure of the sentiment expressed in news articles and social media, offering an additional layer of information that can be harnessed for trading decisions.

FinBERT, a domain-specific language model, has demonstrated promising capabilities in financial sentiment analysis. Fine-tuning FinBERT on relevant datasets can further enhance its performance for specific tasks. In this context, fine-tuning FinBERT using the Financial Phrase Bank dataset aims to improve its ability to interpret and classify financial news headlines related to cryptocurrencies accurately.

The motivation behind this project is to bridge the gap between sentiment analysis and practical trading strategies in the cryptocurrency market. By developing an agent that integrates sentiment scores with technical indicators, we aspire to create a more resilient and adaptive trading strategy capable of navigating the complexities of the cryptocurrency landscape.

**Main Contribution**

The primary contributions of this project are as follows:

1. **Fine-Tuning FinBERT:** We fine-tuned the pre-trained FinBERT model using the Financial Phrase Bank dataset through a cross-validation approach to enhance its performance in financial sentiment analysis.
2. **Development of FINMEMAgent:** We designed and implemented FINMEMAgent, a trading agent that processes financial news, analyzes sentiments, and makes trading decisions based on both sentiment scores and technical indicators.
3. **Comprehensive Trading Simulation:** We conducted an extensive trading simulation using historical data, assessing the agent's performance based on portfolio value, Sharpe ratio, maximum drawdown, and other key metrics.
4. **Performance Analysis:** We evaluated the impact of training epochs on model performance, identifying issues related to overfitting and proposing strategies to mitigate such challenges.
5. **Visualization of Sentiment and Trading Performance:** We generated visualizations to illustrate sentiment trends over time and the relationship between portfolio performance and BTC price movements, including trade action markers.

**System Architecture/Framework**

The system architecture is composed of the following components:

1. **Data Collection and Preprocessing:** Historical price data and news headlines for Bitcoin and Ethereum are collected from JSON files. The data undergoes preprocessing, including date parsing, calculation of price changes, and computation of technical indicators like the Simple Moving Average (SMA).
2. **Sentiment Analysis Model (FinBERT):** FinBERT is fine-tuned using the Financial Phrase Bank dataset through a cross-validation approach to enhance its sentiment analysis capabilities for financial texts.
3. **FINMEMAgent:** This agent processes batches of news headlines, computes sentiment scores using the fine-tuned FinBERT model, and maintains a memory of past data to inform trading decisions.
4. **Trading Decision Logic:** The agent makes trading decisions based on sentiment scores and technical indicators. The decision logic prioritizes buying when positive sentiment is high, and the price is above the SMA and selling under negative sentiment or when certain technical conditions are met.
5. **Trading Simulation:** A simulation runs over historical data, executing the agent's decisions and tracking the portfolio's value over time.
6. **Performance Evaluation:** The performance of the trading strategy is assessed using metrics such as total return, Sharpe ratio, maximum drawdown, and trade profitability.

**Proposed Approaches or Methods**

**Data Preprocessing**

* **Loading Data:** Historical price data and news summaries for Bitcoin and Ethereum were loaded from specified JSON files.
* **Data Frame Creation:** The data was transformed into Pandas DataFrames with date indices to facilitate time-series analysis.
* **Calculating Technical Indicators:** Price changes and Simple Moving Averages (SMA) were computed to serve as technical indicators for trading decisions.
* **Data Merging:** Bitcoin and Ethereum datasets were merged to create a unified dataset, ensuring temporal alignment and consistency.

**Fine-Tuning FinBERT**

* **Dataset Utilization:** The Financial Phrase Bank dataset, comprising sentences annotated for sentiment (positive, neutral, negative), was employed for fine-tuning.
* **Cross-Validation Approach:** Stratified K-Fold Cross-Validation with 3 folds was implemented to ensure robust model performance and mitigate overfitting.
* **Tokenization:** Sentences were tokenized using FinBERT's tokenizer with appropriate truncation and padding to handle varying sentence lengths.
* **Training Parameters:**
  + **Epochs:** Set to 3 per fold to balance learning and prevent overfitting.
  + **Batch Size:** 16 for training and 64 for evaluation.
  + **Learning Rate:** 2e-5.
  + **Weight Decay**: 0.01.
* **Training Process:** The Hugging Face Trainer API facilitated the training process, incorporating evaluation strategies and metric computations.
* **Metrics Computation:** Accuracy, precision, recall, and F1-score were calculated to assess model performance during training and evaluation.

**FINMEMAgent Implementation**

* **Processing News:** FINMEMAgent processes news headlines in batches, leveraging the fine-tuned FinBERT model to obtain sentiment scores.
* **Memory Management:** The agent maintains a memory buffer to store recent data, ensuring that decision-making incorporates relevant historical information.
* **Decision Logic:**
* **Positive Sentiment:** Initiates a BUY action if positive sentiment exceeds a defined threshold and the BTC price is above the SMA.
* **Negative Sentiment:** Initiates a SELL action if negative sentiment surpasses a set threshold or if the BTC price drops significantly below the SMA.
* **Holding:** Maintains the current position if sentiment and technical indicators do not meet the criteria for buying or selling.

**Trading Simulation**

* **Initial Portfolio:** The simulation commences with an initial cash amount (e.g., $10,000) and no cryptocurrency holdings.
* **Execution of Decisions:** The agent's BUY, SELL, or HOLD decisions are executed sequentially, adjusting the portfolio accordingly based on BTC prices.
* **Portfolio Tracking:** The value of the portfolio is continuously tracked, accounting for cash and cryptocurrency holdings.

**Performance Evaluation**

**Metrics Calculated:**

* **Total Return:** Percentage increase in portfolio value over the simulation period.
* **Sharpe Ratio:** Risk-adjusted return measure.
* **Maximum Drawdown:** The largest peak-to-trough decline in portfolio value.

**Visualization:** Graphical representations of sentiment trends over time and portfolio performance relative to BTC price movements were generated to provide intuitive insights.

**Experiments Results and Analysis**

**Data Analysis and Visualization**

* **Price Trends:** The historical price trends of Bitcoin and Ethereum were analyzed using moving averages to identify patterns and potential trading signals.
* **News Volume:** The volume of news headlines varied over the simulation period, with certain periods exhibiting high activity and others being relatively quiet. This variability influenced the sentiment analysis and subsequent trading decisions.

**Model Fine-Tuning Results**

* **Cross-Validation Performance:** Through Stratified K-Fold Cross-Validation, the fine-tuned FinBERT model demonstrated consistent performance across all folds.
* **Label Distribution in Financial PharseBank:**

|  |  |
| --- | --- |
| **Label** | **Count** |
| Negative (0) | 303 |
| Neutral (1) | 1391 |
| Positive (2) | 570 |

**Evaluation Metrics:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Metric** | **Fold 1** | **Fold 2** | **Fold 3** | **Average** |
| Accuracy | 97.48% | 98.54% | 99.34% | 98.45% |
| Precision | 95.79% | 97.17% | 98.81% | 97.26% |
| Recall | 96.36% | 97.81% | 98.66% | 97.61% |
| F1 Score | 96.04% | 97.48% | 98.73% | 97.42% |

**Overfitting Observation:** Limiting the number of training epochs to three per fold prevented overfitting, as evidenced by consistent or decreasing validation loss across epochs without significant increases in training loss.

**Trading Simulation Results**

* **Agent’s Decision:**

|  |  |
| --- | --- |
| **Decision** | **Count** |
| BUY | 3 |
| SELL | 2 |
| HOLD | 83 |

* **Portfolio Performance:**

|  |  |
| --- | --- |
| **Metric** | **Value** |
| Initial Value | $10,000.00 |
| Final Value | $16,974.01 |
| Total Return | 68.74% |
| Sharpe Ratio | 2.67 |
| Maximum Drawdown | -0.08 |

* **Trade Details:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Buy Date** | **Sell Date** | **Buy Price** | **Sell Price** | **Profit($)** |
| 2022-11-30 | 2022-22-16 | 17168.57 | 16647.48 | -521.08 |
| 2022-12-25 | 2023-02-24 | 16841.99 | 23198.13 | 6356.18 |
| 2023-03-12 | N/A | 22163.95 | N/A | N/A |

**Total Profit from Trades:** $5835.06

**Visualization**

**A graph of lines with numbers and letters

Description automatically generated with medium confidence**

Figure : Sentiment Overtime

A graph with numbers and lines

Description automatically generated with medium confidence

Figure : Portfolio Value vs. BTC Price

**Analysis and Interpretation of Results**

* **Profitability:** The agent achieved a total return of 69.74%, demonstrating effective decision-making based on sentiment and technical indicators.
* **Risk-Adjusted Return:** A Sharpe ratio of 2.67 indicates that the returns were achieved with an excellent level of risk, showcasing efficient portfolio management.
* **Drawdowns:** The maximum drawdown of -0.08 signifies periods of decline were minimal, highlighting the agent's ability to mitigate significant losses effectively.
* **Decision Distribution:** The agent predominantly held positions, with a few BUY and SELL actions, reflecting a balanced and strategic trading approach.
* **Impact of Sentiment Analysis:** Incorporating sentiment analysis provided valuable context for trading decisions, enabling the agent to respond to market sentiments reflected in news headlines. This is evident from the profitable trades made based on sentiment-driven signals.

**Discussion of Errors and Open Questions**

**Overfitting in Model Training:**

* **Cause:** While limiting training epochs helped prevent overfitting, the high performance across all folds suggests the model generalizes well on the given dataset. However, there's a possibility that the model might overfit if exposed to more diverse or larger datasets.
* **Effect:** If overfitting were to occur, it could reduce the model's ability to accurately predict sentiments on unseen data, potentially impacting the reliability of trading decisions.

**Data Sparsity:**

* **Issue:** On days with limited or no news headlines, the sentiment analysis was less informative, necessitating reliance on technical indicators alone.
* **Impact:** This could lead to missed opportunities or suboptimal trading decisions during periods of low news activity.

**Market Volatility:**

* **Challenge:** Sudden market shifts not captured by news headlines could result in delayed or inappropriate trading actions by the agent.
* **Effect:** The agent's performance may be adversely affected during unexpected market events, underscoring the need for more dynamic response mechanisms.

**Trade Execution Timing:**

* **Issue:** The simulation assumes immediate execution of trades based on daily sentiment and price data. In real-world scenarios, there might be delays or slippage affecting trade execution.
* **Impact:** This could lead to discrepancies between simulated and actual trading performance.

**Addressing Identified Errors**

* **Limiting Epochs:** Capped the number of training epochs at three to prevent overfitting and ensure better generalization to unseen data.
* **Data Augmentation:** Future work could involve augmenting the dataset with additional examples to provide more diverse training data, enhancing the model's robustness.
* **Adaptive Strategies:** Implementing adaptive strategies to account for days with limited news data, such as weighting technical indicators more heavily or integrating alternative data sources.
* **Real-Time Data Integration**: Incorporating real-time data streams and responsive mechanisms to better handle sudden market volatility and improve decision-making accuracy.
* **Trade Execution Refinement:** Enhancing the simulation to account for trade execution delays and slippage to better mirror real-world trading conditions.

**Conclusions**

The integration of sentiment analysis into cryptocurrency trading strategies demonstrates significant potential. Our fine-tuned FinBERT model, utilized by FINMEMAgent, effectively informed trading decisions that led to a substantial return of 69.74% in the simulated environment. The project's findings indicate that leveraging financial news sentiment can enhance trading strategies by providing contextual insights into market movements.

However, challenges such as overfitting during model fine-tuning and data sparsity in news headlines necessitate careful consideration and mitigation. The observed prevention of overfitting through limited training epochs underscores the importance of balanced training procedures, while data sparsity highlights the need for comprehensive data sources and adaptive trading strategies.

Overall, the project underscores the viability of sentiment-driven trading approaches in the cryptocurrency market, offering a foundation for further exploration and refinement.

**Future Work**

1. **Enhancing Data Sources:** Incorporate additional data sources such as social media platforms (e.g., Twitter) to enrich sentiment analysis and capture a broader spectrum of market sentiments.
2. **Advanced Models:** Explore more sophisticated models or ensemble methods to improve sentiment classification accuracy and robustness.
3. **Risk Management Strategies:** Implement advanced risk management techniques within the agent to better mitigate drawdowns and manage portfolio volatility.
4. **Real-Time Trading Adaptation:** Adapt FINMEMAgent for real-time trading scenarios, enabling live market interactions and immediate response to sentiment shifts.
5. **Multi-Cryptocurrency Support:** Extend the agent's capabilities to support a broader range of cryptocurrencies, enhancing its applicability and market coverage.
6. **Incorporating Additional Indicators:** Integrate more technical indicators (e.g., Relative Strength Index, MACD) to provide a more comprehensive basis for trading decisions.
7. **Trade Execution Refinement:** Improve the simulation to account for real-world trading factors such as slippage, transaction fees, and execution delays.

**Limitations**

1. **Data Limitations**: The reliance on historical price data and news headlines may not fully capture the multifaceted nature of market dynamics and external influences.
2. **Model Generalization:** The model's ability to generalize to entirely new or unseen data is constrained, as evidenced by potential overfitting tendencies if exposed to more diverse datasets.
3. **Simplified Trading Logic:** The decision-making framework is based on a simplified set of rules and may not account for the myriad factors influencing cryptocurrency markets, potentially limiting its effectiveness in complex scenarios.
4. **Latency in Sentiment Processing:** The time required to process large batches of news headlines may introduce latency, affecting the timeliness of trading decisions in fast-moving markets.
5. **Portfolio Diversification:** The simulation focused primarily on Bitcoin and Ethereum, lacking diversification across other cryptocurrencies, which could impact overall portfolio performance and risk exposure.
6. **Assumption of Perfect Trade Execution:** The simulation assumes immediate and exact execution of trades based on the agent's decisions, which may not reflect real-world trading conditions where delays and slippage can occur.

**Lessons Learned**

1. **Importance of Monitoring Training:** Vigilant monitoring of training and validation metrics is essential to prevent overfitting and ensure that the model maintains the ability to generalize effectively.
2. **Integration of Multiple Indicators:** Combining sentiment analysis with technical indicators can enhance trading strategies but requires meticulous calibration to balance the influence of each component.
3. **Complexity of Financial Markets:** Financial markets are influenced by a vast array of factors and developing models that are robust and adaptable to various scenarios is crucial for sustained success.
4. **Data Quality and Quantity:** The quality and quantity of data play a pivotal role in model performance. Ensuring comprehensive and diverse datasets is fundamental to building effective sentiment analysis models.
5. **Adaptive Decision-Making:** Incorporating mechanisms that allow for adaptive decision-making based on real-time data and changing market conditions can significantly enhance trading strategy effectiveness.
6. **Trade Execution Realism:** Simulating more realistic trade execution conditions can provide a more accurate assessment of a trading strategy's performance.

**Acknowledgements**

We extend our sincere gratitude to Dr. Liao for his invaluable guidance and support throughout this project. Additionally, we acknowledge the creators of FinBERT and the providers of the Financial Phrase Bank dataset for their contributions, which were instrumental in the development and fine-tuning of our sentiment analysis model.

**References**

1. **FinBERT Model**: <https://github.com/ProsusAI/finBERT>
2. **Financial Phrase Bank Dataset**: Malo, P., Sinha, A., Korhonen, P., Wallenius, J., & Takala, P. (2014). Good debt or bad debt: Detecting semantic orientations in economic texts. *Journal of the Association for Information Science and Technology*, 65(4), 782-796.
3. **Hugging Face Transformers Library**: <https://huggingface.co/transformers/>
4. **FinBERT Model**  
   ProsusAI. (2021). FinBERT: A Pretrained Financial Language Model. Retrieved from <https://github.com/ProsusAI/finbert>
5. **Backtrader**  
   Backtrader. (n.d.). Backtrader: Python Backtesting Library for Financial Strategies. Retrieved from <https://www.backtrader.com/>
6. **Momentum Backtrader**  
   Mementum. (2021). Momentum: Backtrader Framework for Cryptocurrency Trading. Retrieved from <https://github.com/mementum/backtrader>
7. **Sharpe Ratio**  
   Wikipedia contributors. (2023). Sharpe Ratio. Wikipedia, The Free Encyclopedia. Retrieved from <https://en.wikipedia.org/wiki/Sharpe_ratio>
8. **Crypto Trading Research**  
   Fin.ai. (2025). Crypto Trading Insights: A Comprehensive Guide. Retrieved from <https://coling2025cryptotrading.thefin.ai/>