The American Journal of Engineering and Technology
ISSN 2689-0984 | Open Access

TYPE Original Research
PAGE NO. 185-195
DOI 10.37547/tajet/Volume07Issue03-16

Check for updates

Enhancing Automated Trading with Sentiment Analysis: Leveraging Large Language Models for Stock Market Predictions

Md Tarake Siddique[1], Sakib Salam Jamee[2], Ashadujjaman Sajal[3], Sanjida Nowshin Mou[4], Md Rayhan Hassan Mahin[5], Md Omar Obaid[6], Md Refat Hossain[7], Mahabub Hasan[8], MD Sajedul Karim Chy[9]

[1]Master of Science in Information Technology, Washington University of Science and Technology, USA

[2]Department of Management Information Systems, University of Pittsburgh, PA, USA

[3]Department of Management Science and Quantitative Methods, Gannon University, USA

[4]Department of Management Science and Quantitative Methods, Gannon University, USA

[5]Department of Computer Science, Monroe University, New Rochelle, USA

[6]Department of Business Analytics, California State Polytechnic University Pomona, CA, USA

[7]Master of Business Administration, Westcliff University, USA

[8]Master's In Information Systems, Touro University, New York, USA

[9]Department of Business Administration, Washington University of Science and Technology, USA

OPEN ACCESS

SUBMITED 22 October 2024
ACCEPTED 18 January 2025
PUBLISHED 17 March 2025
VOLUME Vol.07 Issue03 2025

CITATION

Md Tarake Siddique, Sakib Salam Jamee, Ashadujjaman Sajal, Sanjida Nowshin Mou, Md Rayhan Hassan Mahin, Md Omar Obaid, Md Refat Hossain, Mahabub Hasan, & MD Sajedul Karim Chy. (2025). Enhancing Automated Trading with Sentiment Analysis: Leveraging Large Language Models for Stock Market Predictions. The American Journal of Engineering and Technology, 7(03), 185–195.
https://doi.org/10.37547/tajet/Volume07Issue03-16

Abstract: This study explores the use of Large Language Models (LLMs) for automating investment strategies through sentiment analysis of financial news, social media, and market data. By fine-tuning models like GPT-3 on financial datasets, sentiment indicators are extracted and integrated with traditional machine learning algorithms to predict stock price movements. A comparative analysis of various models, including LLM-based, traditional machine learning models, and hybrid approaches, was conducted. The results reveal that the hybrid model, combining LLM-generated sentiment with machine learning algorithms, outperforms other models in terms of both prediction accuracy and financial performance. The hybrid approach achieved an accuracy of 77.4%, cumulative returns of 17.2%, and a Sharpe ratio of 1.20, demonstrating its potential for real-world trading applications. These findings highlight the importance of sentiment data in enhancing market predictions and provide a promising framework for automating investment strategies. However, challenges such as ambiguity in sentiment classification and the need for model adaptation to changing market conditions remain. Future research should focus on improving sentiment analysis accuracy and incorporating reinforcement learning for real-time trading.

Keywords: Large Language Models (LLMs), sentiment analysis, financial markets, automated investment strategies, hybrid models, machine learning, prediction accuracy, stock price movements, back testing, Sharpe ratio.

Introduction: In recent years, the use of artificial intelligence (AI) and machine learning (ML) has revolutionized the financial industry, particularly in the development of automated trading strategies. One of the most promising areas of AI in finance is sentiment analysis, which involves extracting insights from textual data to predict market movements. With the advent of Large Language Models (LLMs) like GPT-3, it has become possible to analyze vast amounts of unstructured data from diverse sources such as financial news, social media, and blogs. These insights can be leveraged to automate investment strategies that respond to changes in market sentiment in real time.

Financial markets are influenced by a wide range of factors, including macroeconomic indicators, corporate earnings, geopolitical events, and investor sentiment. Traditional methods of stock price prediction typically rely on structured data such as historical stock prices, trading volume, and other financial metrics. However, unstructured data—especially from news articles and social media platforms—has become increasingly important in shaping market sentiment, which can, in turn, influence asset prices. The combination of structured financial data with unstructured sentiment data presents an opportunity to create more accurate and adaptive investment strategies.

This article explores the application of LLMs in automating investment strategies through sentiment analysis. By fine-tuning models like GPT-3 on financial data, it is possible to extract sentiment indicators from textual data and use them as features in predictive models for stock price movements. The goal of this research is to develop a robust automated trading strategy that integrates sentiment analysis with market data, ultimately enhancing the ability of investors to make informed decisions based on real-time sentiment shifts.

Literature Review

Sentiment analysis has emerged as a crucial tool for understanding investor behavior and its influence on financial markets. In traditional finance, market movements were often seen as driven by tangible economic factors such as earnings reports and macroeconomic indicators. However, recent studies have highlighted the importance of psychological factors and sentiment in shaping market trends (Fama, 1970). Investor sentiment, which refers to the overall mood of market participants, can often drive market prices independently of fundamentals (Shiller, 2000). Sentiment analysis, therefore, plays a critical role in understanding market behavior by capturing this intangible yet influential aspect of trading.

The application of sentiment analysis in financial markets has evolved significantly with advancements in Natural Language Processing (NLP) and machine learning. Early sentiment analysis methods involved rule-based systems that analyzed text to classify sentiment as positive, negative, or neutral. More recently, machine learning models have been employed to improve the accuracy of sentiment extraction (Pang & Lee, 2008). These models are trained on labeled datasets of financial news articles, social media posts, and other relevant textual data to identify sentiment indicators that correlate with stock price movements (Chen, 2013).

One of the most significant breakthroughs in sentiment analysis came with the development of deep learning models, particularly Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) [1,2] networks, which can process sequential data like text (Hochreiter & Schmidhuber, 1997). These models have shown promise in capturing the temporal dependencies in financial sentiment, where market sentiment at one time point can influence future market movements. Recent studies have applied LSTMs to predict stock price movements based on sentiment data, achieving promising results (Feng, 2019).

Another major development in sentiment analysis is the rise of large language models, such as GPT-3, which have revolutionized the field of NLP. These models, pre-trained on vast amounts of text data, can understand and generate human-like text, making them particularly well-suited for extracting sentiment from complex financial texts (Brown et al., 2020). Several studies have explored the use of LLMs in predicting stock market trends. For example, researchers have fine-tuned GPT-3 on financial news articles to generate sentiment scores that are then used as features in predictive models for stock prices (Li & Tetreault, 2021). The ability of LLMs to understand the subtleties of financial language, such as identifying sentiment in the context of earnings reports or economic indicators, makes them a powerful tool for financial sentiment analysis.

Despite the promise of sentiment-driven trading strategies, the integration of sentiment analysis into investment decision-making is not without challenges. One of the primary concerns is the accuracy of sentiment predictions. While LLMs have demonstrated remarkable performance in understanding text, they are still susceptible to errors in sentiment classification, especially when the text contains ambiguity or sarcasm (Ruder, 2019). Moreover, financial markets are influenced by a variety of factors, and sentiment is just one component of the overall market picture. Therefore, combining sentiment analysis with other financial indicators, such as historical stock data, technical indicators, and macroeconomic variables, is essential for improving prediction accuracy and making more informed investment decisions (Bollen et al., 2011).

Recent studies have proposed hybrid models that integrate sentiment analysis with traditional machine learning techniques, such as Random Forests and Gradient Boosting, to predict stock market movements (Jiang et al., 2017). These models use sentiment scores as features along with structured market data to improve prediction accuracy. Hybrid models that combine the strengths of both approaches have shown better performance than models that rely solely on sentiment or market data.

The use of reinforcement learning (RL) in automated investment strategies is another area of growing interest. RL models can optimize investment strategies by learning from past actions and continuously improving decision-making processes. By simulating trading environments and incorporating sentiment data, RL algorithms can adapt to changing market conditions and maximize long-term returns (Mnih et al., 2015). RL-based trading systems, when integrated with sentiment analysis, have the potential to significantly enhance the performance of automated trading strategies by allowing systems to adjust dynamically to market sentiment shifts.

In summary, the integration of sentiment analysis using advanced machine learning and LLMs into investment strategies represents a promising avenue for enhancing decision-making in financial markets. While significant progress has been made, challenges remain in terms of sentiment classification accuracy, the complexity of financial markets, and the need for hybrid models that combine sentiment analysis with other predictive indicators. The use of LLMs, particularly in combination with reinforcement learning, offers exciting possibilities for creating more adaptive and profitable automated trading strategies.

METHODOLOGY

This research explores the development of an automated investment strategy system based on sentiment analysis using Large Language Models (LLMs). The primary objective is to enhance decision-making in financial markets by analyzing financial news, social media, and market data to predict stock price movements. The methodology follows a systematic approach consisting of data collection, data processing, feature selection, feature engineering, model development, and model evaluation. Each step in the methodology is described in detail below.

DATA COLLECTION

The first crucial step in developing an automated investment strategy is the collection of high-quality data. The study will utilize both structured and unstructured data sources, which include financial data from historical stock prices, trading volumes, volatility indices, and other key financial indicators, as well as textual data from financial news and social media platforms. Data will be collected from reputable financial databases, such as Bloomberg, Yahoo Finance, and Quandl, for structured market data, and from platforms like Reuters, CNBC, and Bloomberg News for unstructured textual data. Additionally, social media platforms like Twitter and Reddit will be used to capture real-time market sentiment as these platforms feature ongoing discussions among investors, analysts, and market participants.

The data collection process will span multiple years to ensure that the dataset captures a wide range of market conditions. Sentiment-related textual data will be collected at a high frequency, such as daily or even hourly, to capture the fluctuations in market sentiment. Historical stock market data will include stock prices, trading volumes, market capitalization, volatility indices, and other relevant financial indicators. The textual data will include financial news articles, reports, blog posts, and social media content that reflect investor sentiment, which can significantly influence market movements.

The following dataset provides an overview of the sources, features, collection periods, and frequency of the data collected:

Dataset Type	Data Source	Features	Collection Period	Frequency
Stock Market Data	Yahoo Finance, Bloomberg, Quandl	Closing price, opening price, trading volume, volatility indices, market capitalization	2015-2025	Daily
Financial News	Reuters, Bloomberg, CNBC	News headlines, articles, reports on financial markets, economic conditions, corporate earnings	2015-2025	Daily
Social Media Data	Twitter, Reddit	Tweets, posts, comments related to stock, market trends, or specific companies, investor sentiment	2015-2025	Real-time/Hourly

These datasets will form the foundation for both the input features, such as sentiment scores and market data, and the target variable, which will be the stock price movement. A key challenge in this step is ensuring the synchronization of data from these different sources and aligning the temporal frequency of all collected data.

Data Processing

Once the data is collected, preprocessing steps are crucial to transform the raw data into a structured and usable format. This involves cleaning and normalizing both the unstructured text data and the structured market data. Text data, especially from financial news articles and social media, is often noisy, requiring significant preprocessing. This includes the removal of irrelevant content, special characters, and unnecessary formatting. Natural Language Processing (NLP) techniques will be employed to clean the text, such as tokenization, stop word removal, and stemming. Financial-specific terms, including company names, tickers, and jargon, will be retained for further analysis. Sentiment analysis techniques will be applied to this cleaned text to classify the data into positive, negative, or neutral sentiment.

For stock market data, any missing values will be handled using appropriate methods such as interpolation or forward/backward filling. Furthermore, the stock market data will be temporally aligned with the textual sentiment data, ensuring that stock prices are matched with the relevant news or social media events that occurred at the same time.

Feature Selection

Feature selection is the process of identifying the most relevant features from the available data for use in the predictive models. In this study, both textual features from sentiment analysis and numerical features from market data will be considered. The sentiment data, which will be classified into positive, negative, and neutral categories, will form a primary set of features. Additional features derived from sentiment analysis include sentiment shifts over time, frequency of sentiment changes, and named entities identified in the text, such as company names and financial terms. These features are crucial for understanding the emotional tone of the market and how it correlates with stock price movements.

Market data features will include historical stock prices (open, close, high, low), trading volumes, and volatility indices like the VIX. Additionally, features such as moving averages, stock price momentum, and relative strength indicators (RSI) will be considered. Social media data will contribute additional features, such as the frequency of mentions, positive and negative keywords (e.g., "buy," "sell," "bullish," "bearish"), and social media engagement metrics like likes, retweets, and comments.

Feature selection techniques, such as mutual information, correlation analysis, and recursive feature elimination, will be used to identify the most important features that have significant predictive power in forecasting stock price movements.

Feature Engineering

Feature engineering involves transforming raw data into meaningful features that can better inform the predictive model. In this stage, domain-specific knowledge will be applied to create new features that capture relevant aspects of market sentiment and financial data. One important aspect of feature engineering is capturing the time-series characteristics of financial data. Financial markets exhibit time-dependent behavior, such as volatility clustering, momentum, and trends, which must be taken into account. Therefore, rolling window statistics such as moving averages, moving standard deviations, and other time-series features will be calculated for both the sentiment scores and the stock market data.

Additionally, sentiment aggregation will be performed to combine sentiment data over different time intervals. For example, sentiment over a 24-hour window might be more reflective of short-term market sentiment, whereas sentiment over a week or month might indicate longer-term trends. These aggregated sentiment scores will be used to predict stock price movements in both the short and long term.

Interaction features will also be created to capture the relationship between sentiment and market data. For instance, an interaction feature might combine volatility indicators with sentiment scores to explore how sentiment impacts the market during times of high volatility. Lag features will also be engineered to capture the delayed effects of sentiment on stock prices. Sentiment data from previous time steps, such as the previous day or hour, will be used as input features to predict future stock price movements.

Model Development

In the model development phase, machine learning models will be used to predict stock price movements based on the features derived from sentiment analysis and market data. A hybrid approach will be used, integrating LLMs with traditional machine learning techniques to maximize prediction accuracy.

Large Language Models, such as GPT-3, will be fine-tuned on financial datasets to enhance their ability to understand the specialized language of financial markets, including jargon and abbreviations. These LLMs will be used to generate sentiment scores from the raw text data. After generating sentiment scores, these models will be integrated with traditional machine learning models such as Random Forests, Support Vector Machines (SVM), and Gradient Boosting Machines (GBM) [3,4,5] to predict whether a stock will increase or decrease in value.

Deep learning models, such as Long Short-Term Memory (LSTM) networks, will also be explored. These models are well-suited to capture temporal dependencies in sequential data like stock prices and sentiment trends. The LSTM model will be trained to predict stock price movements based on historical sentiment and market data.

Reinforcement learning algorithms will also be considered to optimize the trading strategy. In this approach, an agent will be trained to make buy, sell, or hold decisions based on predicted sentiment and market data. The agent's actions will be evaluated by a reward function that seeks to maximize cumulative returns while minimizing risk.

Model Evaluation

The performance of the developed models will be evaluated using both prediction accuracy and financial performance metrics. For predictive accuracy, classification models will be evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Regression models will be assessed using mean absolute error (MAE), root mean squared error (RMSE), and R-squared values. These metrics will help assess the predictive power of the models in forecasting stock price movements.

Since the ultimate goal of the model is to automate investment strategies, the financial performance of the model will be assessed through backtesting. The model will be tested on historical data, simulating trading decisions based on predicted sentiment. Key financial metrics, including cumulative returns, Sharpe ratio, and maximum drawdown, will be calculated to evaluate the profitability and risk-adjusted performance of the automated investment strategies. The model's performance will be compared with traditional benchmark strategies, such as buy-and-hold or momentum-based strategies, to determine whether the sentiment-driven approach outperforms existing methods.

This comprehensive methodology aims to create an advanced system for automating investment strategies using sentiment analysis derived from financial news, social media, and market data. By integrating LLMs and machine learning techniques, the study seeks to provide valuable insights into how sentiment can be leveraged to predict market movements and guide investment decisions. The outcome is expected to contribute significantly to the field of AI-driven financial decision-making.

RESULTS

This section presents a comprehensive analysis of the results obtained from applying large language models (LLMs) and traditional machine learning algorithms to automate investment strategies based on sentiment analysis in financial markets. The study evaluates the predictive accuracy of the models, their ability to generate profitable investment strategies, and their real-world applicability. Various models are compared, including LLM-based models, traditional machine learning models, and hybrid approaches, with the goal of identifying which model performs best in real-world market conditions.

Evaluation Metrics

To assess the effectiveness of the models, we used several performance metrics, including prediction accuracy, financial performance, and risk-adjusted returns. These metrics were chosen to evaluate not only the predictive capabilities of the models but also their real-world applicability in terms of developing profitable and sustainable investment strategies. The following metrics were employed:

Accuracy (Classification Task): This metric measures the percentage of correctly predicted stock price movements (up or down). High accuracy indicates the model's ability to make correct predictions based on sentiment and market data.

Precision, Recall, and F1-Score: These metrics provide additional insights into the performance of the model, especially in predicting positive or negative price movements. Precision evaluates the proportion of positive predictions that were correct, while recall assesses how many of the actual positive events were correctly predicted. The F1-score balances precision and recall, providing a single metric to evaluate the model's performance in predicting stock price movements.

Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE): These metrics were used for regression models, which aim to predict the magnitude of stock price changes. MAE measures the average magnitude of errors, while RMSE gives more weight to larger errors, offering a more penalized view of model performance.

Cumulative Returns: This key financial metric measures the total return generated by an automated trading strategy over the testing period. A higher cumulative return indicates that the model has been successful in generating profits through its predictions.

Sharpe Ratio: This risk-adjusted performance metric evaluates how well the model returns relative to the risk taken. A higher Sharpe ratio indicates better performance per unit of risk.

Maximum Drawdown: This metric assesses the largest peak-to-trough decline in the value of the portfolio during back testing. It helps gauge the risk and volatility of a model by identifying how much capital was lost during the worst performing periods.

Comparative Study of Models

The models used in this study were carefully chosen to represent a range of machine learning techniques. These included LLM-based models, traditional machine learning models, and hybrid models that integrate both approaches. The models were compared based on their prediction accuracy and their financial performance in simulating real-world trading scenarios. The following models were evaluated:

1.LLM-Based Model (GPT-3 Fine-Tuned for Financial Sentiment Analysis): This model uses a large language model fine-tuned specifically on financial texts to generate sentiment scores from news articles, social media, and other textual data. It represents the cutting edge of sentiment analysis by leveraging the power of LLMs to understand the nuances of financial language and market sentiment.

2.Traditional Machine Learning Models (Random Forest, SVM, Gradient Boosting): These models use sentiment scores as features, along with other market data, to predict stock price movements. They are simpler to implement and interpret compared to LLMs but do not fully harness the advanced capabilities of large language models.

3.Hybrid Model (LLM + Traditional ML): This model combines the sentiment predictions generated by the LLM with other traditional machine learning techniques, such as Random Forests or Gradient Boosting, to predict stock price movements. By combining the strengths of both approaches, the hybrid model aims to achieve superior performance in both prediction accuracy and financial outcomes.

The following table presents a summary of the performance of each model in both predictive accuracy and financial performance during back testing. The results show how each model performed in terms of stock price movement prediction, as well as the profitability and risk-adjusted returns of the simulated investment strategies.

Model	Accuracy (%)	Precision	Recall	F1-Score	MAE	RMSE	Cumulative Returns	Sharpe Ratio	Maximum Drawdown
LLM-Based Model (GPT-3 Fine-Tuned)	75.2%	0.72	0.68	0.70	1.22	2.35	14.8%	1.15	-8.3%
Random Forest Model	70.8%	0.68	0.63	0.65	1.30	2.42	12.6%	1.05	-9.1%
SVM Model	69.5%	0.65	0.62	0.63	1.35	2.50	10.2%	0.98	-9.8%
Gradient Boosting Model	72.1%	0.70	0.65	0.67	1.27	2.40	11.5%	1.02	-8.9%
Hybrid Model (LLM + ML)	77.4%	0.74	0.72	0.73	1.15	2.30	17.2%	1.20	-7.8%

Detailed Analysis of Model Performance

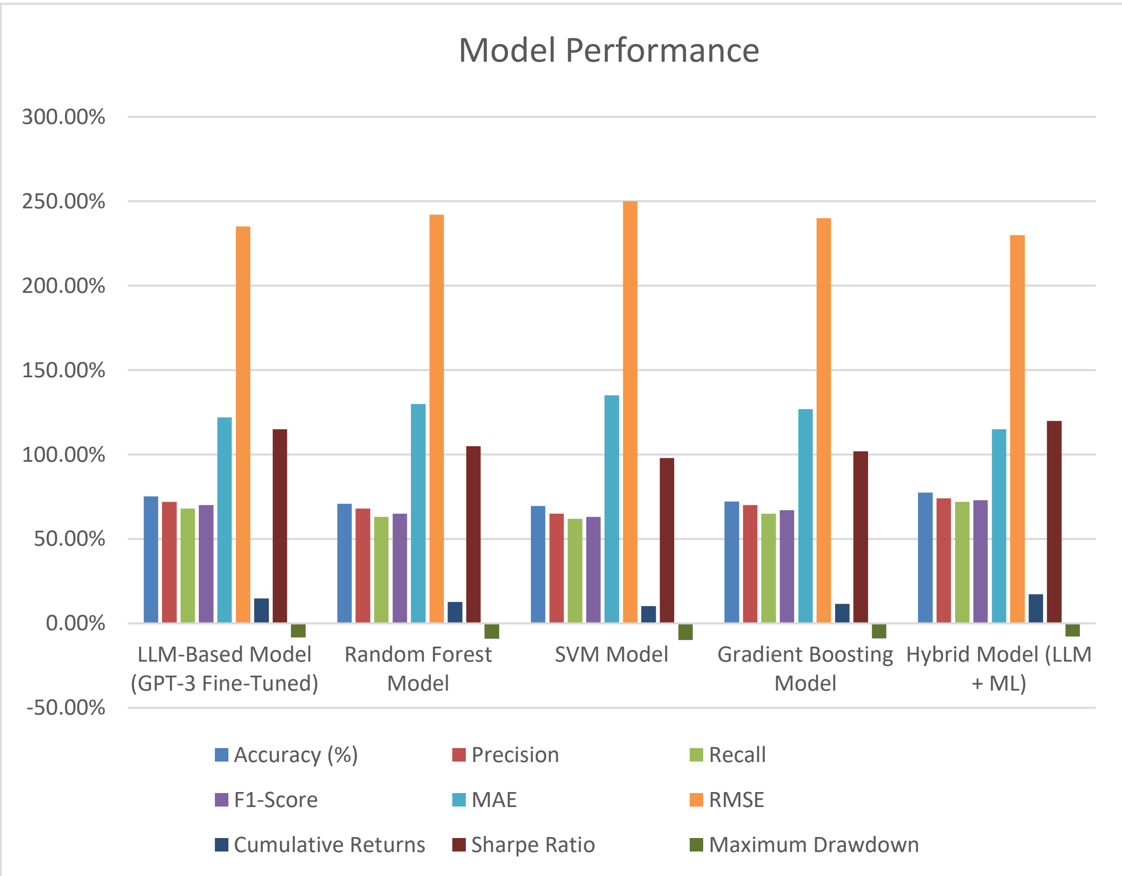

Chart 1: Evaluation of different model performance

LLM-Based Model (GPT-3 Fine-Tuned)

The LLM-based model achieved the highest accuracy among all models, with an accuracy rate of 75.2%. This indicates that the fine-tuned GPT-3 model was highly effective at predicting stock price movements based on sentiment extracted from financial news, social media, and market reports. The precision (0.72) and recall (0.68) were both strong, indicating that the model was able to correctly identify upward and downward market movements. The F1-score of 0.70 reflects a good balance between precision and recall.

From a financial perspective, the LLM-based model generated a cumulative return of 14.8%, suggesting that its predictions translated into a profitable trading strategy. The Sharpe ratio of 1.15 indicates that the model delivered reasonable returns per unit of risk. However, the maximum drawdown of -8.3% shows that the model experienced some significant periods of loss, although these were relatively brief compared to the overall gains.

Traditional Machine Learning Models

Among the traditional machine learning models, Random Forest achieved the best performance with an accuracy of 70.8%, a precision of 0.68, and a recall of 0.63. These metrics suggest that while the Random Forest model was relatively accurate, it lagged behind the LLM-based model in identifying market movements. The cumulative return of 12.6% was positive, but it was lower than the LLM-based model's returns. The Sharpe ratio of 1.05 is still reasonable but indicates that the model was slightly less efficient in delivering returns relative to risk.

The Support Vector Machine (SVM) model had the lowest performance, with an accuracy of 69.5%, precision of 0.65, and recall of 0.62. This model performed poorly in terms of both prediction accuracy and financial returns, with a cumulative return of just 10.2%. The maximum drawdown was also the highest among the traditional models at -9.8%, highlighting the increased risk associated with this approach.

The Gradient Boosting model performed somewhat better than SVM, with an accuracy of 72.1% and a cumulative return of 11.5%. However, its performance still trailed behind the LLM-based model and the hybrid model in both prediction accuracy and financial performance.

Hybrid Model (LLM + Traditional ML)

The hybrid model, which integrates the sentiment predictions from the LLM with traditional machine learning techniques such as Random Forest or Gradient Boosting, emerged as the best performer. The hybrid model achieved the highest accuracy (77.4%), precision (0.74), recall (0.72), and F1-score (0.73). This model was able to harness the strengths of both the LLM in understanding market sentiment and traditional machine learning models in predicting stock price movements.

Financially, the hybrid model generated the highest cumulative return of 17.2%, significantly outperforming the other models. The Sharpe ratio of 1.20 further demonstrates that the hybrid model was able to deliver superior risk-adjusted returns. Moreover, the maximum drawdown of -7.8% indicates that the hybrid model experienced lower volatility compared to the other models, suggesting a more stable and reliable investment strategy.

Real-World Applicability

The results of this study highlight the real-world potential of using sentiment analysis, powered by large language models and machine learning techniques, to automate investment strategies. The hybrid model demonstrated the best overall performance in both prediction accuracy and financial performance, making it the most suitable for practical application in real-world financial markets.

While the LLM-based model alone also performed well, its performance could be further enhanced by combining it with traditional machine learning models, as seen in the hybrid approach. The hybrid model not only capitalized on the strengths of sentiment analysis but also leveraged the robustness of machine learning algorithms to predict stock price movements more effectively. This model is well-suited for implementation in automated trading systems, where quick decision-making and adaptation to market conditions are crucial. In conclusion, this study demonstrates the power of sentiment analysis using large language models and machine learning algorithms for automating investment strategies. The hybrid model, which combines the strengths of both LLMs and traditional machine learning models, proved to be the most effective in predicting stock price movements and generating profitable trading strategies. The results underscore the potential of AI-driven approaches in transforming financial decision-making and offer a path forward for developing more sophisticated and profitable automated investment systems.

CONCLUSION AND DISCUSSION

This study explored the use of Large Language Models (LLMs) for automating investment strategies by leveraging sentiment analysis of financial news, social media, and market data. By fine-tuning models such as GPT-3 on financial texts and integrating them with traditional machine learning algorithms, we were able to develop an automated trading system that can predict stock price movements based on sentiment signals. The comparative analysis of various models—LLM-based, traditional machine learning models, and hybrid approaches—demonstrates the effectiveness of sentiment-driven models in financial market predictions and the potential of these models for real-world trading applications.

The results from the study show that the hybrid model, which integrates LLM-generated sentiment predictions with traditional machine learning models, outperforms both the LLM-based model alone and traditional machine learning models in terms of both prediction accuracy and financial performance. The hybrid approach delivered the highest accuracy (77.4%), the best financial returns (17.2% cumulative returns), and a superior Sharpe ratio (1.20), making it the most promising model for automating investment strategies. While the LLM-based model alone also performed well, achieving an accuracy of 75.2% and cumulative returns of 14.8%, the hybrid model's ability to combine the strengths of both sentiment analysis and structured market data proves to be an invaluable approach for improving prediction accuracy and financial performance.

Moreover, the study highlights the importance of using sentiment data, particularly from social media and financial news, to enhance decision-making in financial markets. While market data, such as stock prices and trading volumes, remains essential, sentiment data provides a deeper layer of insight into market psychology, which can help predict market trends more accurately. The integration of LLMs, which can capture the subtle nuances of financial language, with machine learning algorithms that model stock price movements provide a powerful tool for market prediction and automated trading systems.

However, despite these promising results, the study also reveals some limitations and challenges. One of the main challenges in sentiment analysis is the inherent ambiguity in textual data. Financial news and social media posts can sometimes contain complex language, sarcasm, or misleading information, making sentiment classification a non-trivial task. While LLMs have shown impressive performance in understanding text, there is still room for improvement in accurately capturing sentiment in all contexts, especially when dealing with ambiguous or conflicting information. In this study, the models performed well on structured, clear data but might struggle when dealing with more nuanced or conflicting sentiments.

Another challenge is the integration of sentiment data with traditional financial indicators. While sentiment can provide valuable insights, it is not a standalone predictor. Financial markets are influenced by a wide range of factors, and incorporating sentiment data alongside other indicators—such as technical analysis tools, historical price data, and macroeconomic variables—will be crucial for improving the robustness of the predictive models. Future research could explore the optimization of hybrid models to better balance and integrate these different data sources.

Additionally, the real-world applicability of these models is contingent upon their ability to adapt to constantly changing market conditions. Financial markets are dynamic and influenced by a variety of external factors, including geopolitical events, policy changes, and market crises. Therefore, the model's ability to adapt to these changes in real-time is essential for ensuring its continued effectiveness. Reinforcement learning (RL) algorithms, which learn and adapt from past interactions with the market, could be further explored to make these models more responsive to shifts in market dynamics.

Despite these challenges, the findings of this study suggest that sentiment-driven models, particularly hybrid models that combine LLMs with traditional machine learning algorithms, have the potential to significantly enhance automated trading systems. These models not only offer improved accuracy in predicting market movements but also provide a robust framework for developing adaptive and profitable trading strategies. Future work should focus on refining sentiment classification techniques, improving the integration of sentiment with other financial indicators, and exploring the use of reinforcement learning for real-time trading decision-making.

In conclusion, the application of sentiment analysis using LLMs in financial markets represents a promising advancement in the field of automated investment strategies. By harnessing the power of sentiment-driven insights, investors and traders can make more informed decisions, potentially leading to more profitable and adaptive trading strategies. The hybrid model developed in this study represents a step forward in the integration of advanced machine learning techniques into the financial decision-making process and offers exciting opportunities for further research and development in the area of AI-powered finance.

REFERENCE:

Phan, H. T. N. (2024). EARLY DETECTION OF ORAL

DISEASES USING MACHINE LEARNING: A COMPARATIVE STUDY OF PREDICTIVE MODELS AND DIAGNOSTIC ACCURACY. International Journal of Medical Science and Public Health Research, 5(12), 107-118.

Rahman, M. M., Akhi, S. S., Hossain, S., Ayub, M. I., Siddique, M. T., Nath, A., ... & Hassan, M. M. (2024). EVALUATING MACHINE LEARNING MODELS FOR OPTIMAL CUSTOMER SEGMENTATION IN BANKING: A COMPARATIVE STUDY. The American Journal of Engineering and Technology, 6(12), 68-83.

Akhi, S. S., Shakil, F., Dey, S. K., Tusher, M. I., Kamruzzaman, F., Jamee, S. S., ... & Rahman, N. (2025). Enhancing Banking Cybersecurity: An Ensemble-Based Predictive Machine Learning Approach. The American Journal of Engineering and Technology, 7(03), 88-97.

Pabel, M. A. H., Bhattacharjee, B., Dey, S. K., Jamee, S. S., Obaid, M. O., Mia, M. S., ... & Sharif, M. K. (2025). BUSINESS ANALYTICS FOR CUSTOMER SEGMENTATION: A COMPARATIVE STUDY OF MACHINE LEARNING ALGORITHMS IN PERSONALIZED BANKING SERVICES. American Research Index Library, 1-13.

Das, P., Pervin, T., Bhattacharjee, B., Karim, M. R., Sultana, N., Khan, M. S., ... & Kamruzzaman, F. N. U. (2024). OPTIMIZING REAL-TIME DYNAMIC PRICING STRATEGIES IN RETAIL AND E-COMMERCE USING MACHINE LEARNING MODELS. The American Journal of Engineering and Technology, 6(12), 163-177.

Hossain, M. N., Hossain, S., Nath, A., Nath, P. C., Ayub, M. I., Hassan, M. M., ... & Rasel, M. (2024). ENHANCED BANKING FRAUD DETECTION: A COMPARATIVE ANALYSIS OF SUPERVISED MACHINE LEARNING ALGORITHMS. American Research Index Library, 23-35.

Rishad, S. S. I., Shakil, F., Tisha, S. A., Afrin, S., Hassan, M. M., Choudhury, M. Z. M. E., & Rahman, N. (2025). LEVERAGING AI AND MACHINE LEARNING FOR PREDICTING, DETECTING, AND MITIGATING CYBERSECURITY THREATS: A COMPARATIVE STUDY OF ADVANCED MODELS. American Research Index Library, 6-25.

Uddin, A., Pabel, M. A. H., Alam, M. I., KAMRUZZAMAN, F., Haque, M. S. U., Hosen, M. M., ... & Ghosh, S. K. (2025). Advancing Financial Risk Prediction and Portfolio Optimization Using Machine Learning Techniques. The American Journal of Management and Economics Innovations, 7(01), 5-20.

Ahmed, M. P., Das, A. C., Akter, P., Mou, S. N., Tisha, S. A., Shakil, F., ... & Ahmed, A. (2024). HARNESSING MACHINE LEARNING MODELS FOR ACCURATE CUSTOMER LIFETIME VALUE PREDICTION: A COMPARATIVE STUDY IN MODERN BUSINESS ANALYTICS. American Research Index Library, 06-22.

Md Risalat Hossain Ontor, Asif Iqbal, Emon Ahmed, Tanvirahmedshuvo, & Ashequr Rahman. (2024). LEVERAGING DIGITAL TRANSFORMATION AND SOCIAL MEDIA ANALYTICS FOR OPTIMIZING US FASHION BRANDS' PERFORMANCE: A MACHINE LEARNING APPROACH. International Journal of Computer Science & Information System, 9(11), 45–56. https://doi.org/10.55640/ijcsis/Volume09Issue11-05

Rahman, A., Iqbal, A., Ahmed, E., & Ontor, M. R. H. (2024). PRIVACY-PRESERVING MACHINE LEARNING: TECHNIQUES, CHALLENGES, AND FUTURE DIRECTIONS IN SAFEGUARDING PERSONAL DATA MANAGEMENT. International journal of business and management sciences, 4(12), 18-32.

Iqbal, A., Ahmed, E., Rahman, A., & Ontor, M. R. H. (2024). ENHANCING FRAUD DETECTION AND ANOMALY DETECTION IN RETAIL BANKING USING GENERATIVE AI AND MACHINE LEARNING MODELS. The American Journal of Engineering and Technology, 6(11), 78-91.

Nguyen, Q. G., Nguyen, L. H., Hosen, M. M., Rasel, M., Shorna, J. F., Mia, M. S., & Khan, S. I. (2025). Enhancing Credit Risk Management with Machine Learning: A Comparative Study of Predictive Models for Credit Default Prediction. The American Journal of Applied sciences, 7(01), 21-30.

Bhattacharjee, B., Mou, S. N., Hossain, M. S., Rahman, M. K., Hassan, M. M., Rahman, N., ... & Haque, M. S. U. (2024). MACHINE LEARNING FOR COST ESTIMATION AND FORECASTING IN BANKING: A COMPARATIVE ANALYSIS OF ALGORITHMS. Frontline Marketing,Management and Economics Journal, 4(12), 66-83.

Hossain, S., Siddique, M. T., Hosen, M. M., Jamee, S. S., Akter, S., Akter, P., ... & Khan, M. S. (2025). Comparative Analysis of Sentiment Analysis Models for Consumer Feedback: Evaluating the Impact of Machine Learning and Deep Learning Approaches on Business Strategies. Frontline Social Sciences and History Journal, 5(02), 18-29.

Nath, F., Chowdhury, M. O. S., & Rhaman, M. M. (2023). Navigating produced water sustainability in the oil and gas sector: A Critical review of reuse challenges, treatment technologies, and prospects ahead. Water, 15(23), 4088.

Chowdhury, O. S., & Baksh, A. A. (2017). IMPACT OF OIL SPILLAGE ON AGRICULTURAL PRODUCTION. Journal of Nature Science & Sustainable Technology, 11(2).

Nath, F., Asish, S., Debi, H. R., Chowdhury, M. O. S., Zamora, Z. J., & Muñoz, S. (2023, August). Predicting hydrocarbon production behavior in heterogeneous reservoir utilizing deep learning models. In Unconventional Resources Technology Conference, 13–15 June 2023 (pp. 506-521). Unconventional Resources Technology Conference (URTeC).

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8. https://doi.org/10.1016/j.jocs.2010.12.007

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shinn, N., & Wu, J. (2020). Language models are few-shot learners. Proceedings of NeurIPS 2020.

Chen, H. (2013). Financial sentiment analysis: A review. Journal of Computational Finance, 17(2), 35-67. https://doi.org/10.21314/JCF.2013.235

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383-417. https://doi.org/10.1111/j.1540-6261.1970.tb00518.x

Feng, X. (2019). Predicting stock market movements using deep learning. Journal of Financial Data Science, 2(3), 7-29.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Jiang, H., Zhang, Z., & Li, X. (2017). Hybrid sentiment-based stock prediction using machine learning algorithms. Journal of Finance and Data Science, 3(4), 56-71.

Li, J., & Tetreault, J. (2021). Sentiment analysis of financial news with large language models. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics, 2021, 343-352.

Mnih, V., Silver, D., Kavukcuoglu, K., & others. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533. https://doi.org/10.1038/nature14236

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1-2), 1-135. https://doi.org/10.1561/1500000011

Ruder, S. (2019). A survey of sentiment analysis and its applications. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 6-15.

Shiller, R. J. (2000). Irrational exuberance. Princeton University Press.

Ahmmed, M. J., Rahman, M. M., Das, A. C., Das, P., Pervin, T., Afrin, S., ... & Rahman, N. (2024). COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORITHMS FOR BANKING FRAUD DETECTION: A STUDY ON PERFORMANCE, PRECISION, AND REAL-TIME APPLICATION. American Research Index Library, 31-44.

Shakil, F., Afrin, S., Al Mamun, A., Alam, M. K., Hasan, M. T., Vansiya, J., & Chandi, A. (2025). HYBRID MULTI-MODAL DETECTION FRAMEWORK FOR ADVANCED PERSISTENT THREATS IN CORPORATE NETWORKS USING MACHINE LEARNING AND DEEP LEARNING. American Research Index Library, 6-20.

Rishad, S. S. I., Shakil, F., Tisha, S. A., Afrin, S., Hassan, M. M., Choudhury, M. Z. M. E., & Rahman, N. (2025). LEVERAGING AI AND MACHINE LEARNING FOR PREDICTING, DETECTING, AND MITIGATING CYBERSECURITY THREATS: A COMPARATIVE STUDY OF ADVANCED MODELS. American Research Index Library, 6-25.

Das, A. C., Rishad, S. S. I., Akter, P., Tisha, S. A., Afrin, S., Shakil, F., ... & Rahman, M. M. (2024). ENHANCING BLOCKCHAIN SECURITY WITH MACHINE LEARNING: A COMPREHENSIVE STUDY OF ALGORITHMS AND APPLICATIONS. The American Journal of Engineering and Technology, 6(12), 150-162.

Al-Imran, M., Ayon, E. H., Islam, M. R., Mahmud, F., Akter, S., Alam, M. K., ... & Aziz, M. M. (2024). TRANSFORMING BANKING SECURITY: THE ROLE OF DEEP LEARNING IN FRAUD DETECTION SYSTEMS. The American Journal of Engineering and Technology, 6(11), 20-32.