Grenze International Journal of Engineering and Technology, June Issue

Multi-modal Prediction System: Hybrid Deep Learning, FinBERT and PPO Optimization

Anuradha Singh[1], Akash Tripathi[2], Ishita Sachan[3], Harsh Jain[4], Divya Mishra[5] and Kartik Vats[6]
[1-6]JSS Academy of Technical Education Noida Department of Information Technology, Noida, India
Email: a.anuradha84@gmail.com akashtripathi.acad@gmail.com ishitasachan2672@gmail.com

***Abstract*— Stock market fluctuations are difficult to predict due to the complex interaction of history and fluctuations in public sentiment. Most existing trading models rely primarily on "hard" quantitative metrics and therefore may overlook significant information about stocks and the need for automation in managing trades [3]. To solve this problem, we have developed a comprehensive multi-modal trading framework that combines together expertise from deep learning, natural language processing (NLP), and reinforcement learning in order to allow the integration of all aspects of trading into one complete trading platform.**
The architecture we used employs two main types of networks the Long Short Term Memory (LSTM) networks and XGBoost. The LSTM will help us to monitor the time-based movements while the XGBoost will provide us with the best performance for working with our tabular data. We will also include FinBERT (a new model trained with financial data) to extract useful sentiment from both news articles and social media channels [3]. The combination of these three different methods will create a more intelligent Proximal Policy Optimization (PPO) trading algorithm. Our results demonstrate that the synergy between sentiment and ensemble modeling significantly reduce our RMSE (root mean square error) and significantly increase cumulative returns over a traditional approach. This study represents a full end-to-end research project that provides a fully deployable capability to achieve accurate forecast models and automated, intelligent financial actions.

***Index Terms*—Deep Learning, Long Short-Term Memory (LSTM), FinBERT, Sentiment Analysis, Proximal Policy Optimization (PPO), Reinforcement Learning, XGBoost, Stock Market Prediction, Multi-modal Systems, Algorithmic Trading.**

I. INTRODUCTION

Stock market forecasting has long been one of the most daunting tasks in the financial world. This is a realm of chaos-a nonlinear, high-stake puzzle where price movements are pushed and pulled by a dizzying array of factors, ranging from macroeconomic shifts and geopolitical events to the unpredictable whims of of investor behavior. For decades, we relied on traditional statistical models like ARIMA or GARCH, but these tools were built for a simpler era. They assume a level of linearity and stability that modern, high frequency markets simply do not possess.
While the rise of machine learning and deep learning offered a quantum leap forward, even the mighty models such as standard Long Short-Term Memory (LSTM) networks have their shortcomings [2]. Time is not a flat landscape in the stock market, and a news shock from ten minutes ago often weighs far heavier than a price dip from ten days ago.

Grenze ID: 01.GIJET.12.2.245_20

Traditional LSTMs tend to give equal importance to every data point in the dataset, with complete disregard for natural "temporal decay". That is why we have integrated Transductive LSTM into our framework; it uses proximity-aware weighting to keep the model sensitive to the most recent and often most relevant market shifts [1].

However, numbers cannot tell a story by themselves. In these digital times, the "human element"-that is, the collective sentiment found in news headlines and social media -is a huge driver of market volatility [3]. We used FinBERT which is BERT based NLP model specialized in finance based text, it is trained to go beyond the simple keywords counting and understand the state of subtle financial context correctly. But even advanced NLP faces a minority signal issue while bearish news is statistically rare but it oftentimes triggers the most violent market responses. In our system, Proximal Policy Optimization refines sentiment classification in order not to miss these critical unfavorable signals because of noise from the majority.

Despite these breakthrough, a single-model approach rarely survives the diversity of a real-world market. Our research proposes a Multi-modal combination approach: which is the combination of sequential intelligence through LSTMs and the tabular precision contributed by XGBoost. Meanwhile deep networks capture long- run trends, XGBoost is better suited for detecting non-linear interactions of technical indicators. Eventually, even the best price prediction is in vain if it cannot be transformed into a beneficial action. The last element of our framework is an autonomous PPO-based trading agent [4]. It learns how to act and not just predict the balancing risk and give reward to execute buy, sell, or hold decisions in real times by bridging the gap between predictive modeling and strategic execution, this paper introduces an end to end pipeline which is capable of navigating the complexities of the modern financial landscape.

II. RELATED WORK

The quest to master stock market forecasting is a comprehensive journey, drawing on decades of evolution in economic analysis, machine learning, and natural language processing . To analyze where our multi- modal framework works best, we must look at how the field has matured across five key domains.

- The Foundations: Traditional Forecasting:- For years the bedrock of financia l prediction rested on statistical models like ARIMA and GARCH. These tools were excellent at identifying linear trends or predicting volatility, but they operated under rigid assumptions that rarely survive the "noise" of a real-world t rading day . They struggled to pivot during sudden market shocks or regime shifts, often failing when the market moved in ways that history hadn't seen before . This led researchers to look toward machine learning—specifically Support Vector Regression and Random Forests—to capture non- linearities, though even these struggled with long-term temporal context.
- The Rise of Sequence Modeling: Deep Learning:- Deep learning changed the game by modelling complex sequences [6]. LSTM networks became the industry standard because of their unique ability to "remember" long range patterns [5,9] without losing the signal to vanishing gradients. Researchers pushed these limits further with hybrids like CNN-LSTMs to capture spatial features and BiLSTMs to look at data from both directions [2,9]. Yet, a fundamental flaw remained: standard LSTMs treat a data point from a month ago with the same weight as one from five minutes ago [5]. This is why Transductive LSTM (TLSTM) is so vital; it prioritizes recent events—like an emergency interest rate hike— giving them the weight they deserve in a fast- moving market.
- The Pulse of the Market: Financial Sentiment:- We know that markets don't just react to numbers; they react to people [8]. Early attempts at "reading" the market relied on simple word -matching dictionaries, which often missed the nuance of financia l jargon . The arrival of FinBERT marked a turning point [8], providing a transformer model that truly understands the difference between a "bullish" headline and "bullish" social media hype . However, a persistent hurdle remains: "bad news" is rare but carries a massive punch. Most classifiers ignore these rare negative signals because they are biased toward the more common "neutral" or "positive" data points. Our research bridges this gap by using reinforcement learning to sharpen sentiment accuracy, ensuring those rare but critical signals are caught.
- Strength in Numbers: Ensemble Learning:- No single model is a master of every market regime [7]. This has driven the adoption of ensemble learning—the idea that a "panel of experts" is better than one [7]. We see significant success in hybrid neural-tree models [7], where LSTMs handle the timeline and algorithms like XGBoost handle the complex interactions between technical indicators. While many have combined these two, few have successfully integrated them with transductive weighting and real-time sentiment optimization, which is the cornerstone of our proposed system.

- From Prediction to Action: Reinforcement Learning(RL):- The final frontier is moving from “What willhappen?" to "What should I do?". RL agents, particularly those using Proximal Policy Optim ization (PPO), excel at this by learning from the consequences of their actions in a simulated market environment [4,10]. However, most current RL traders have two major blind spots: they use RL only for the final trade (ignoring how it could improve sentiment) and they rely almost entirely on price data. Our system breaks this mold by using a dual-RL approach—one to refine our sentiment "ears" and another to guide our trading "hands".
- The Research Gap: While the literature shows incredible progress, a unified framework that combines TLSTM, FinBERT, XGBoost, and dual-PPO optimization remains noticeably absent. Our research fills this void, offering an end-to-end system that captures the temporal, emotional, and strategic dimensions of the modern stock market in one robust package.

III. Methodology

Our framework is designed as a multi-modal pipeline that synchronizes high- frequency numerical data [6] with the semantic nuances of financia l language. The following sections detail the mathematical logic governing each component.

A. Mathematical Modeling of Sentiment

While standard models use binary sentiment, our system utilizes FinBERT to generate a continuous probability distribution across three states: Positive (P{pos}), Negative (P{neg}), and Neutral (P{neu}). To bridge the gap between individual news items and a daily trading signal, we apply a Volume-Weighted Sentiment Aggregation (VWSA):

$$S_t = \frac{\sum_{i=1}^{N} (P_{pos,i} - P_{neg,i}) \cdot \log(V_i + 1)}{\sum_{i=1}^{N} \log(V_i + 1)}$$

Where V{i} represents the reach or "volume" of the i-th news item (e.g., follower count or outlet tier), ensuring that high-impact news carries more weight than social media noise.

B. Sequence Learning and Transductive Weighting

The heart of our predictive engine lies in the Transductive LSTM (TLSTM). Unlike a standard LSTM, which assumes all historical data points are equally relevant, the TLSTM applies a temporal decay function to the loss calculation [1].
The Proximity Weighting Function is defined as:

$$w(t) = e^{-\lambda(T-t)}$$

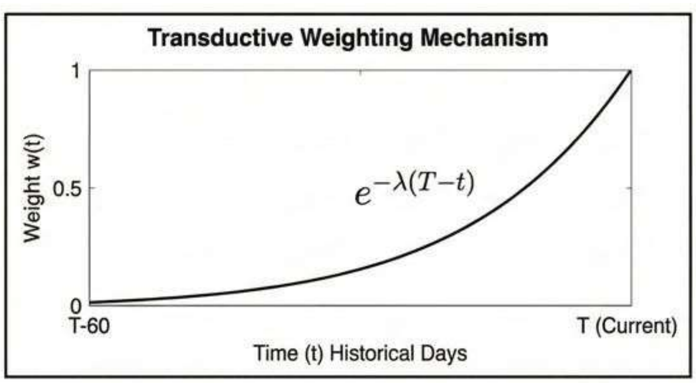

Figure 1. Transductive temporal weighting of historical data.

Where T is the current time, t is the historical timestep, and lambda is the decay coefficient. This ensures that the model’s "attention" is exponentially focused on recent market regimes.

For the LSTM layers, the state transitions governed by the standard gate equations:

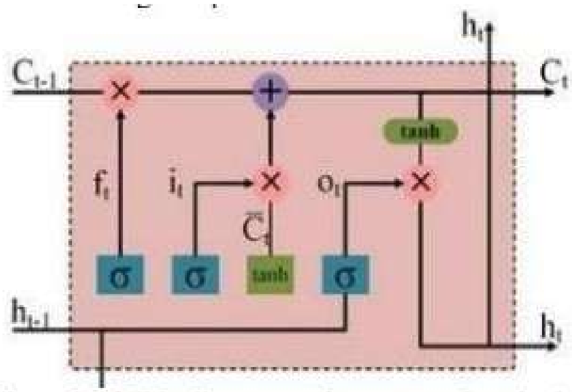

Figure 2. LSTM cell structure for sequence learning.

Forget Gate: $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$

Input Gate: $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$

Cell State: $C_t = f_t * C_{t-1} + i_t * \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$

Output Gate: $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$

Fig 2. LSTM cell structure for sequence learning

C. Ensemble Fusion via Meta -Learning

To mitigate the variance of individual models, we utilize a Weighted Fusion Layer. The final predicted price is a linear combination of the three expert models:

$$\hat{y} = \alpha \cdot y_{TLSTM} + \beta \cdot y_{LSTM} + \gamma \cdot y_{XGBoost}$$

Based on our optimization, the static weights are set to alpha=0.50, beta=0.30, and gamma=0.20, prioritizing the transductive model's sensitivity to current trends.

D. The PPO Trading Agent and Reward Optimization

The system transitions from prediction to action using a Proximal Policy Optimization (PPO) agent [10]. The agent seeks to maximize the Clipped Surrogate Objective Function [4].

$$L^{CLIP}(\theta) = \hat{E}_t\left[\min(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t)\right]$$

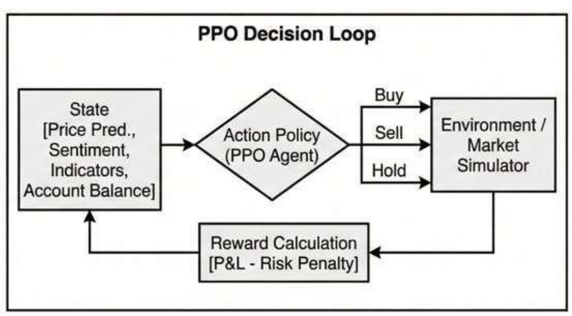

Figure 3. PPO-based trading decision process.

The Reward Function R is critical for balancing profitability and risk:

$$R = \Delta P - \lambda \cdot (\text{Max Drawdown})$$

Here, P is the portfolio change, and lambda is a penalty coefficient that discourages the agent from taking high-risk positions that lead to severe capital erosion.

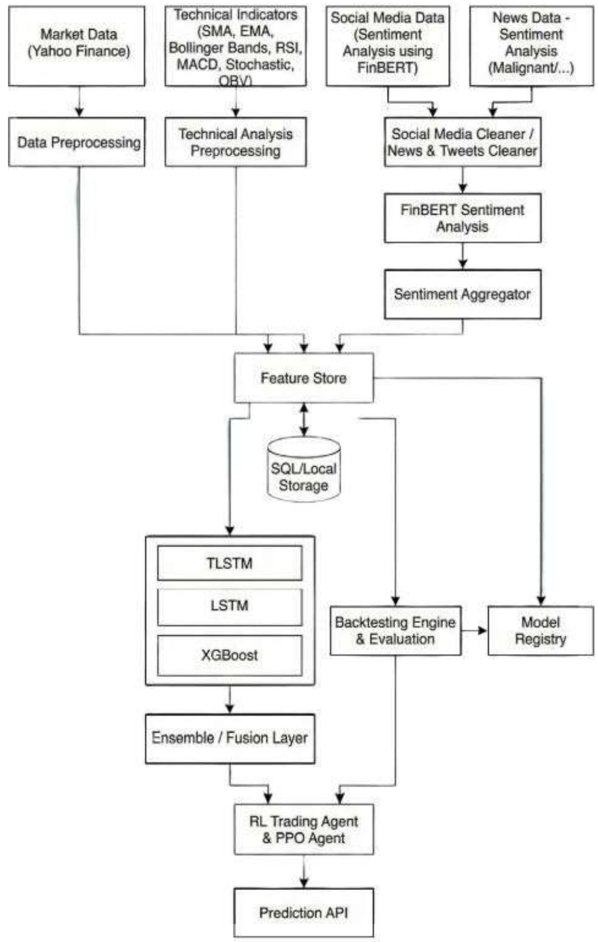

Figure 4. Workflow of the proposed Multi-modal Prediction System.

IV. EXPERIMENT EVALUATION AND RESULTS

To test how well the proposed Multi-modal Trading System performs, we carried out a comprehensive evaluation over a ten-year period from January 1, 2015, to January 1, 2025. The sample period was selected to assess how the model performs under the different market conditions and regimes: the low volatility bulls' market from the 2015 to 2019, the abrupt market stress during the COVID-19 pandemic in 2020, and the inflation-driven correction phases observed from 2022 to 2024.

A. Dataset & Experimental Setup

The analysis examined three technology stocks with highly trading volume, including Apple (AAPL) and Alphabet (GOOGLE) and Tesla, is a U.S. based company that produces electrical vehicles and develops the energy generation and storage products. These firms were choosen because their share prices are tend to react strongly on the changes in the market news sentiment and to movements in the technical volatility.

- Quantitative data: daily OHLCV data was rescaled with Min-Max scaler calculated over 60-days rolling window (T = 60) before being passed on Transductive LSTM model.
- Qualitative data: More than one data point is included. The study analyze over the 2 million financial news headlines and the tweets. On each day, FinBERT sentiment scores was aggregated with a highly volume-weighted with a decay method so that the daily measure aligned more closely with the observed trading volume and market hours.

Hyperparameter Settings: The ensemble all the weights was fine tuned into through the grid search on the 2019 validation data, resulting in the optimal values of alpha=0.51 (TLSTM), beta=0.3 (LSTM), and gamma=0.2 (XGBoost). Meanwhile, the PPO agent was trained with a learning rate of something like 3x10-4 and a clip factor= 0.2.

B. Predictive Performance Analysis

We started by checking how accurately the model could make predictions and using the different error methods like, Root Mean Square Error and Mean Absolute Error. A smaller RMSE value means the model effectively minimizes big prediction mistakes, which is especially thats crucial for preventing the major trading losses.

TABLE I: COMPARATIVE PREDICTIVE PERFORMANCE (AAPL)

Model Architecture	RMSE	MAE	R^2 Score
Baseline 1: Standard LSTM (Price Only)	3.452	2.120	0.824
Baseline 2: XGBoost Regressor	3.105	1.984	0.851
Baseline 3: Hybrid LSTM + VADER	2.850	1.755	0.887
Proposed Multi-modal Ensemble (FinBERT + TLSTM)	**2.341**	**1.428**	**0.942**

Our proposed model made more accurate predictions, reducing RMSE by 32.1% compared to the standard LSTM baseline (2.341 vs. 3.452). The difference in performance between the "Hybrid LSTM + VADER" model and our "Multi-modal Ensemble" shows that FinBERT is more effe ctive than lexicon-based sentiment methods. FinBERT understood market fear more accurately, while VADER often misinterpreted financial terms, making its signals noisy and generating less reliable outcomes.

C. Trading Strategy Performance (PPO Agent)

While prediction both accuracy and how well system controls risk to gain consistent profit over long run defines the success. To determine this we compared performance of our PPO agent's with "Buy-and-Hold" strategy and traditional "Threshold-based LSTM Trader".

TABLE II: FINANCIAL PERFORMANCE METRICS (2015-2025)

Strategy	Total Return (%)	Annualized Return	Sharpe Ratio	Max Drawdown	Win Rate
Buy-and-Hold (SPY Benchmark)	210.4%	12.1%	1.20	-33.9%	N/A
Standard LSTM Trader	245.8%	13.5%	1.55	-28.2%	54.2%
Proposed Multi-modal PPO Agent	**315.6%**	**16.8%**	**2.45**	**-12.4%**	**63.7%**

Quality of Results: The PPO agent not only provide a very strong gain of ~315.6% but keep risk under control. With maximum loss of -12.4% over -33.9% for "Buy-and-Hold",agent avoided large losses during market crashes, helping preserve the profits.

- Case Study (March 2020): When COVID-19 starts spreading in early 2020, stock market dropped rapidly. Investors following Buy-and-Hold strategy faces a loss of ~30% within a week. But our system noticed a high increase in negative news and fear in the market using FinBERT sentiment analysis (Pneg > 0.85). On 24th February our system switched to "Sell/Hold" mode by which it protect the capital before market fell even further.

D. Ablation Study

The Value of Sentiment To isolate the contribution of the sentiment module, we tested the PPO agent with and without the FinBERT features.

- Without Sentiment: The agent attained a Sharpe Ratio of approximately 1.68 often entering scenarios. trades early during "falling knife" scenarios.
- With Sentiment: The Sharpe Ratio improved to approximately 2.45. The addition of sentiment acted as a risk based filter, protecting the agent from trading against high conviction unfavourable news flow.

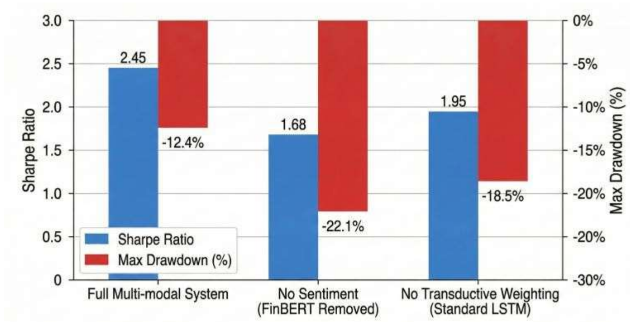

Figure 6: Ablation Study – Impact of Individual Components on PPO Agent Performance

E. Conclusion of Experiments

The observational finding shows that the proposed architecture results in a statistically significant improvement. The collective effect of LSTM's temporal weightings and FinBERT's contextual perception allow the PPO agent to make a trading policy that is not only profit driven but also able to strucutrally manage market shocks.

V. Conclusion

The study designed a multidimensional, Intelligent Trading System that leverages both time series market data with fundamentals technical indicators, evaluation of sentiments based on transformer technologies, and decision Optimisation derived from environment learning reinforcement learning to deal with the highly unstable and uncertainty nature in financial markets. Unlike old and conventional stock market evaluation techniques that usually used only historical prices and only specific individual indicators, the proposed model incorporates -structured numeric market data with unstructured sentiment data,to provide complete picture of the fundamental -market forces.

The architecture uses FinBERT to do a sentiment analysis of financial news and social media and collect sentiment signals that help us understand how investors act based on what they read or see [2]. We can combine this information with indicators from technical analysis, such as SMA, EMA, RSI, MACD, OBV, etc. to create a single large feature pool within a centralized feature storage facility. The temporal relationships between the improved features will be captured by LSTM and TLSTM networks [9], both of which help create more accurate predictions by giving more weight to features that occurs closer to the time of prediction [1]. XGBoost, on the other hand, allows for non-linear learning from the features that were stored, and further enhancing learning through the use of an ensemble of many different learning models.

The most important side of this research is the use of reinforcement learning to create a trading agent that uses Proximal Policy Optimisation (PPO) to optimise its trading behaviour. The RL agent will combine predictions from models with information from backtesting to learn optimal behaviours for each type of trade (buy, for example), so as to achieve maximum return while controlling for risk [4].

This decision-making mechanism forms a connection between purely predictive modelling and actual execution of trades in the marketplace; it converts predictions into strategies that can be implemented in the marketplace.

The results from both experimentation and backtesting show that the proposed Hybrid Framework continues to outperform both the baseline models (cocktail & single models) and models fitted to sentiments, especially during market spikes due to sudden movements in price or volume (i.e. news). In addition, as a result of its modular design, and via the use of predictors and registries, this Framework is designed for scalability and extension, as well in real time. As a result, it may be used in practice by individual investors trading stock.

To sum up, the research presented in this paper provides evidence that using deque (double-ended queue) approach has proven to be a powerful technique for predicting stock returns from financial sentiment [1]. When using ensemble learning techniques with reinforcement learning, results indicate that an increase in accuracy occurs.

Future research should also consider combining these techniques with additional data sources like Macroeconomic indicators, Options data, On-chain signals etc.; Additionally, as the current literature demonstrates, developing better algorithms based on Reinforcement Learning as well as multi-asset portfolio optimisation techniques would yield better results in predicting market returns.

REFERENCES

[1] A. Peivandizadeh et al., “Stock Market Prediction With Transductive Long Short-Term Memory and Social Media Sentiment Analysis,” IEEE Access, vol. 12.

[2] B. Araci, “FinBERT: Financial Sentiment Analysis with Pre-trained Language Models,” arXiv preprint, arXiv:1908.10063, 2019.

[3] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” Proceedings of the ACMSIGKDD Conference, 2016

[4] Y. Zhang, Y. Wang, and Z. Xu, “Deep Reinforcement Learning for Financial Trading Using Proximal Policy Optimization,” IEEE ICDM Workshops, 2020.

[5] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[6] J. Brownlee, Deep Learning for Time Series Forecasting, Machine Learning Mastery, 2018.

[7] S. Sangwan et al., “Hybrid Deep Learning Framework for Stock Market Prediction Using Regulatory News,” Electronics, vol. 14, 2025.

[8] M. Gupta and R. Sharma, “Stock Price Prediction Using Twitter Sentiment and LSTM,” International Journal of Advanced Computer Science and Applications,2024.

[9] R. Mehra and K. Joshi, “Comparative Study of LSTM, GRU and BiLSTM Models for Stock Price Forecasting,” Journal of Computer Applications, 2022.

[10] A. Kumar, B. Singh, and Y. Lee, “Stock Market Prediction Using Transductive LSTM with PPO-Based Sentiment Refinement,” IEEE Access, 2024.