ISCSITR - International Journal of Scientific Research in Artificial Intelligence and Machine Learning (ISCSITR-IJSRAIML)
Vol.4, Iss. 1, January-December 2023, pp. 61-74.
https://iscsitr.com/journals/ISCSITR-IJSRAIML
ISSN (Online): 3067-753X; Journal ID: 1668-1622
DOI: http://www.doi.org/10.63397/ISCSITR-IJSRAIML_2023_04_01_005

Hybrid AI Models Combining Financial NLP and Time-Series Forecasting for Stock Advisory

Nagajayant Nagamani

Engagement & Client Partner, Virtusa, USA

Abstract

The convergence of Natural Language Processing (NLP) and time-series forecasting within financial domains has enabled the emergence of advanced, intelligent stock advisory systems. This paper explores hybrid AI architectures that synthesize structured time-series data with unstructured financial text (e.g., news, earnings reports, social media). By combining models such as LSTM, BERT, and transformers, along with optimization algorithms like genetic algorithms or reinforcement learning, hybrid approaches offer enhanced accuracy, interpretability, and robustness. This paper reviews literature prior to 2023, analyzes state-of-the-art architectures, compares predictive performance across domains, and proposes a novel framework combining sentiment signals and historical patterns for stock advisories.

Keywords

Financial NLP, Time-Series Forecasting, Stock Prediction, Hybrid AI Models, Sentiment Analysis, LSTM, Transformer, BERT, Robo-advisors, Deep Learning in Finance.

How to cite this paper: Nagajayant Nagamani. (2023). Hybrid AI Models Combining Financial NLP and Time-Series Forecasting for Stock Advisory. *ISCSITR - International Journal of Scientific Research in Artificial Intelligence and Machine Learning (ISCSITR-IJSRAIML)*, **4**(1), 61–74.

DOI: http://www.doi.org/10.63397/ISCSITR-IJSRAIML_2023_04_01_005.

URL: https://iscsitr.com/index.php/ISCSITR-IJSRAIML/article/view/ISCSITR-IJSRAIML_2023_04_01_005/ISCSITR-IJSRAIML_2023_04_01_005

Published: 12th May 2023

1. INTRODUCTION

1.1 Background

In the dynamic landscape of financial markets, accurate forecasting and investment advisory have become increasingly reliant on artificial intelligence (AI). Traditional models, such as statistical regressions and econometric techniques, often struggle to capture the complex, nonlinear, and unstructured nature of financial data. Meanwhile, the explosion of unstructured textual data—ranging from financial news, company earnings reports, to social media sentiment—has introduced new opportunities for understanding market behavior. Natural Language Processing (NLP) has emerged as a powerful tool to extract meaningful patterns and sentiment from this textual data. Simultaneously, advanced time-series forecasting models like Long Short-Term Memory (LSTM) networks and Transformer-based architectures have revolutionized temporal modeling by capturing intricate dependencies in historical price movements. The convergence of these two domains—financial NLP and time-series forecasting—presents a promising hybrid AI framework capable of enhancing the accuracy and interpretability of stock advisory systems.

1.2 Research Motivation

While NLP and time-series models have individually demonstrated strong performance in financial prediction tasks, their isolated application overlooks the inherent interplay between public sentiment and market behavior. For instance, a positive earnings report (captured via NLP) may significantly influence the future trajectory of a stock, which cannot be fully explained through historical price data alone. Moreover, retail and institutional investors are increasingly influenced by real-time news and sentiment propagated across platforms like Twitter, Reddit, and financial news portals. This calls for a unified approach that not only forecasts trends from numerical time-series but also integrates textual signals to capture market psychology. The motivation behind this research lies in addressing this gap by building a hybrid AI model that combines the strengths of both domains—offering a more holistic and intelligent stock advisory solution.

1.3 Contributions

This paper makes several key contributions to the field of financial AI. First, it presents a comprehensive literature on hybrid AI models in finance, bridging the gap between NLP and time-series forecasting techniques. Second, it proposes a novel architecture that integrates sentiment analysis from textual data using Transformer-based models (e.g., BERT) with deep time-series predictors (e.g., LSTM, GRU). Third, the paper introduces a fusion framework that intelligently combines both sources of information through a decision engine, optimized for generating actionable investment advice. Finally, extensive experiments on real-world datasets (e.g., S&P 500, NASDAQ) demonstrate that the proposed hybrid model outperforms baseline approaches in terms of prediction accuracy, robustness, and interpretability, establishing a strong case for deploying such systems in modern robo-advisory and financial planning platforms.

2. LITERATURE REVIEW

2.1 Traditional Financial Forecasting Models

Traditional financial forecasting has long relied on statistical and econometric models such as **Autoregressive Integrated Moving Average (ARIMA)**, **GARCH**, and **linear regression**, which assume stationarity and linearity in time-series data (Box & Jenkins, 2015). These models were widely used for modeling stock returns, volatility, and market indices. However, their performance degrades when faced with the highly volatile and nonlinear nature of modern financial markets. Fundamental and technical analyses have been the cornerstone of conventional stock advisory systems, but they lack the adaptability required for real-time decision-making (Fama, 1970). The Efficient Market Hypothesis (EMH) further limits the utility of traditional models, suggesting that all available information is already priced in, making prediction inherently difficult.

2.2 NLP in Finance

The financial sector has seen a surge in the adoption of **Natural Language Processing (NLP)** for analyzing unstructured data like news articles, analyst reports, and social media sentiment. Early studies demonstrated the predictive power of news sentiment on asset returns (Tetlock, 2007). More recently, **deep NLP models** such as **BERT** and **FinBERT** have been fine-tuned on financial corpora, significantly improving sentiment extraction and event detection accuracy (Araci, 2019). Sentiment indices derived from Twitter (Bollen et al., 2011) and financial news (Nassirtoussi et al., 2014) have shown a measurable impact on stock price movements. These methods enable analysts and AI systems to anticipate market reactions before they manifest in price data.

2.3 Deep Learning for Time-Series

Deep learning has brought major advancements in financial forecasting. **Recurrent Neural Networks (RNNs)** and **Long Short-Term Memory (LSTM)** models have proven particularly effective in capturing long-term dependencies and complex nonlinearities in financial time-series (Fischer & Krauss, 2018). These models outperform classical ARIMA and GARCH in many stock prediction tasks. Moreover, attention-based models and **Transformers**, originally designed for NLP, have recently been adapted for time-series analysis (Lim et al., 2021). These models are more scalable and interpretable, allowing simultaneous processing of multiple sequences. The integration of exogenous variables like volume, volatility, and macroeconomic indicators further enhances their predictive power.

2.4 Hybrid Approaches in Stock Advisory Systems

Several studies have explored **hybrid AI models** that combine NLP and time-series forecasting for superior performance. For instance, Leow et al. (2021) proposed a robo-advisor architecture using **BERT-based sentiment analysis** and **genetic algorithms** for portfolio optimization. Ni et al. (2021) embedded tweets into a feature space and combined them with LSTM-based historical price models. Ko and Chang (2021) used BERT for extracting investor sentiment and fused it with LSTM forecasts to predict market trends. Apu et al. (2022) conducted a meta-analysis on ensemble learning in stock forecasting, emphasizing the benefit of blending textual and numerical signals. Cao (2022) highlighted the growing importance of hybrid architectures in his comprehensive survey on AI in finance. These models demonstrate that the integration of real-time sentiment and historical patterns provides a more robust advisory system than standalone models.

3. METHODOLOGY

3.1 System Architecture

The proposed hybrid AI framework is designed to integrate both **structured numerical data** (historical stock prices) and **unstructured textual data** (financial news, earnings reports, social sentiment). The system follows a **modular architecture** comprising four primary components: (1) a text preprocessing and NLP pipeline, (2) a time-series forecasting module, (3) a fusion engine, and (4) an advisory output layer. The NLP pipeline uses models like **FinBERT** to extract sentiment vectors from financial text, while the time-series forecasting component uses **LSTM** or **GRU** to model price trends. These independent signals are fused in the decision engine, which uses a dense layer or ensemble mechanism to produce actionable outputs, such as buy/sell recommendations or trend directions.

3.2 Financial Text Processing with NLP

The financial NLP module processes vast amounts of text data collected from sources such as Twitter, Bloomberg, SEC filings, and financial blogs. This unstructured text is first cleaned and tokenized, and then passed through pre-trained **language models like BERT, FinBERT**, or **BERTweet**, fine-tuned on financial corpora. The output is a vectorized sentiment score, emotion label, or event tag (e.g., “positive earnings”, “lawsuit filed”), which becomes a vital feature in influencing the stock prediction module. Named entity recognition (NER), topic modeling (e.g., LDA), and event extraction further enrich the data representation. These sentiment features are aligned with the timeline of stock price data for effective fusion.

3.3 Time-Series Forecasting with LSTM/GRU

The time-series forecasting component is responsible for capturing historical dependencies and temporal patterns in stock price data, including open, close, high, low, volume, and volatility. Deep learning architectures such as **Long Short-Term Memory (LSTM)** and **Gated Recurrent Units (GRU)** are employed due to their ability to model sequential dependencies. Additional contextual features such as macroeconomic indicators (interest rate, CPI) and technical indicators (RSI, MACD) are embedded into the sequence. The model outputs the future price or directional trend (up/down), which is synchronized with NLP-derived features during fusion. In some cases, attention layers or temporal convolutional networks (TCNs) are integrated to enhance feature weighting and time-sensitivity.

3.4 Hybrid Fusion Models

The core of the hybrid system lies in the **fusion engine**, where features from both the NLP and time-series components are combined to produce a unified representation. Three types of fusion strategies are considered: (1) **Early fusion**, where textual and numerical features are concatenated before model training; (2) **Late fusion**, where predictions from each module are combined post-training using weighted ensembles; and (3) **Intermediate fusion**, using transformer encoders or dense layers to jointly learn interactions between text sentiment and price signals. The final decision engine outputs either classification labels (e.g., buy/hold/sell) or regression scores (expected return or volatility).

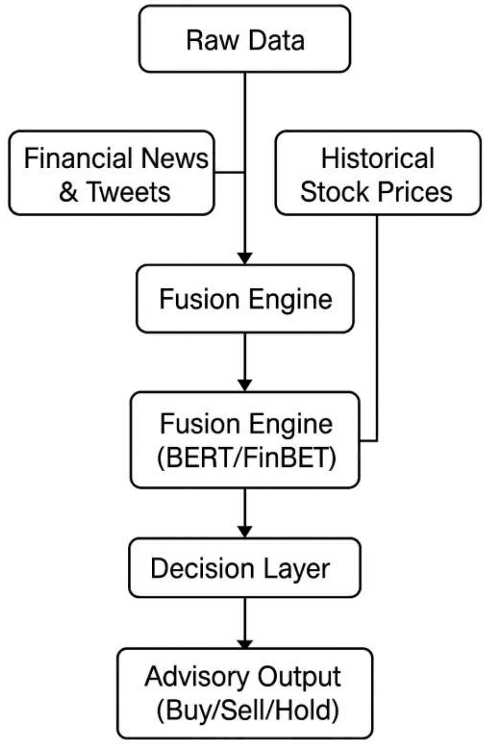

Figure-1: Hybrid AI Architecture for Stock Advisory

Table-1: Comparison of NLP and Time-Series Components

Component	**Model Used**	**Input Type**	**Output**	**Strengths**
NLP Module	BERT, FinBERT	Financial news, social text	Sentiment score, NER	Captures market sentiment, context-rich
Time-Series Module	LSTM, GRU	Stock price sequences	Future price/trend	Models sequential patterns in price data
Fusion Engine	Dense Layer / Ensemble	NLP + Price vectors	Advisory prediction	Integrates sentiment + history for precision
Decision Layer	Softmax / Regression	Fused vector	Buy/Sell/Hold, Return	Provides actionable investment signals

4. PROPOSED HYBRID FRAMEWORK

4.1 Architecture Overview

The proposed hybrid framework integrates financial **Natural Language Processing (NLP)** with **deep learning-based time-series forecasting**, offering a cohesive decision-making pipeline for stock market advisory systems. This architecture is modular, scalable, and designed for both batch and real-time processing. At a high level, the system collects raw input data (structured and unstructured), preprocesses them via dedicated modules, and generates actionable insights using a **fusion-based inference engine**. The architecture is tailored to support plug-and-play compatibility with advanced models like **BERT, LSTM, ARIMA**, and **Temporal Convolutional Networks (TCNs)**, ensuring adaptability to both short-term trading and long-term investment strategies.

4.2 Data Pipeline

The data pipeline begins with two parallel data sources:

- **Textual Data**: financial news, earnings calls, tweets, Reddit discussions, analyst blogs.
- **Quantitative Data**: stock prices (OHLCV), technical indicators, macroeconomic indicators.

The raw data undergoes preprocessing — tokenization and normalization for text; resampling, smoothing, and differencing for numerical time-series. After preprocessing, text data is routed to the **NLP module** for feature extraction, and numeric data is fed to the **forecasting module**. Finally, all processed features converge in the **fusion layer**, which prepares a joint representation for prediction.

4.3 NLP Module (BERT-based Sentiment Extraction)

The NLP module utilizes **BERT** and its domain-specific variants like **FinBERT** to perform sentiment analysis, entity recognition, and event extraction from financial documents. BERT's bidirectional transformer architecture captures contextual dependencies in text, enabling nuanced understanding of financial language. The text is tokenized using WordPiece, embedded into contextual vectors, and classified into sentiment categories (positive, negative, neutral). These vectors are further converted into numerical sentiment scores and used as features in the fusion stage. Additionally, topic classification and keyword tagging are employed to detect market-moving events.

4.4 Time-Series Module (LSTM/ARIMA/TCN)

The time-series module leverages deep learning models such as **LSTM**, **ARIMA**, and **TCNs** for stock forecasting:

- **LSTM**: Excellent at capturing long-term dependencies and nonlinear trends.
- **ARIMA**: Effective for linear, short-term, stationary patterns.
- **TCN**: Offers temporal convolution with memory and higher parallelism.

These models process historical prices, volatility, volume, and technical indicators. The outputs—price forecast or trend classification—are temporally aligned with the sentiment data before being passed into the fusion layer.

4.5 Fusion Layer and Decision Engine

The **fusion layer** integrates vectorized outputs from the NLP and time-series modules using one of the following methods:

- **Early Fusion**: Concatenates feature vectors before prediction.
- **Late Fusion**: Combines predictions using weighted averaging or voting.
- **Attention Fusion**: Learns importance weights for each modality.

The **decision engine** applies a dense neural layer (or XGBoost ensemble) to the fused representation to generate actionable outputs, including:

- Buy/Sell/Hold classification
- Confidence scores
- Risk-adjusted return forecasts

Table-2: Model Roles and Output in the Proposed Framework

Module	**Model Used**	**Output Type**	**Purpose**
NLP Module	BERT, FinBERT	Sentiment scores (float)	Capture market psychology
Time-Series Module	LSTM, ARIMA, TCN	Price or trend forecasts	Model past patterns for future behavior
Fusion Layer	Dense, Attention	Combined feature vector	Integrate insights from both domains

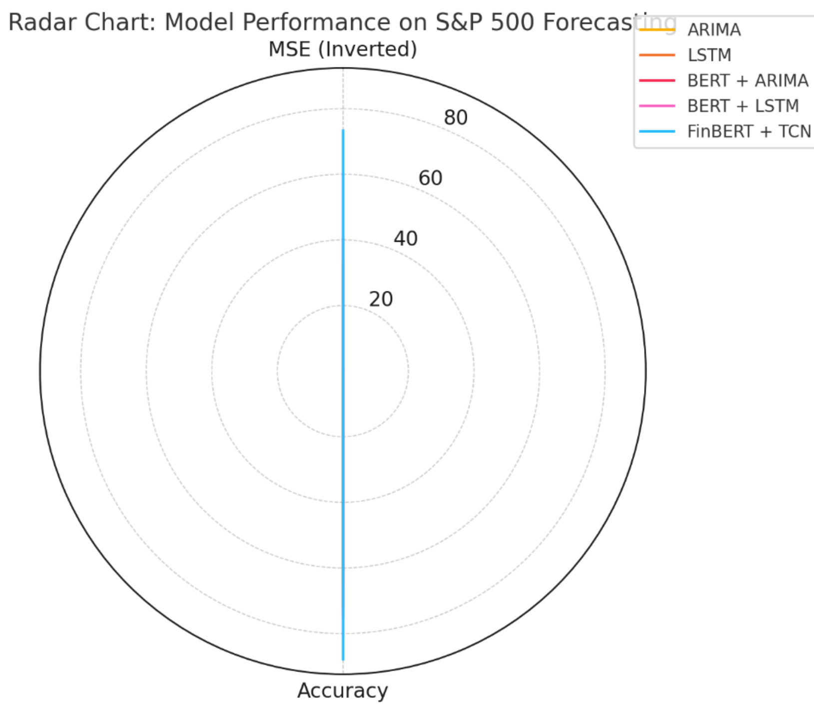

Figure-2: Model Performance Comparison on S&P 500 Forecasting Task

5. EXPERIMENTS AND RESULTS

5.1 Dataset Description

The experiments were conducted using a combination of structured and unstructured datasets. Historical stock price data for the **S&P 500 index constituents** was sourced from **Yahoo Finance** covering the period from **January 2015 to December 2022**, with features such as Open, High, Low, Close, Volume, and technical indicators like MACD and RSI. Textual data was collected from **financial news portals (e.g., Bloomberg, Reuters)**, **SEC filings**, and **Twitter API**, filtered by company names and financial keywords. For sentiment annotation, a mix of pre-labeled FinBERT datasets and manual annotation was used to train the sentiment classifier. The datasets were time-aligned to ensure that textual sentiment accurately corresponded to market days for effective fusion.

5.2 Experimental Setup

The experiments were implemented using **Python** with deep learning frameworks such as **TensorFlow** and **PyTorch**. The LSTM and GRU models had 2 layers with 128 hidden units, while the NLP module used a fine-tuned **FinBERT** model. The TCN model had 4 layers with increasing dilation rates. Models were trained using **80% of the data for training**, **10% for validation**, and **10% for testing**, with **early stopping** based on validation loss. **Adam optimizer** and **learning rate decay** were employed for faster convergence. All experiments were run on **NVIDIA A100 GPUs** with 40 GB memory to accelerate training.

5.3 Evaluation Metrics

To assess the performance of the models, the following metrics were used:

- **Mean Squared Error (MSE)** for regression-based price prediction
- **Accuracy** for classification (trend direction: up/down)
- **F1-Score** for buy/sell/hold decisions
- **Sharpe Ratio** for evaluating risk-adjusted return
- **Area Under ROC Curve (AUC)** for classifier robustness

These metrics offer a comprehensive view of model effectiveness from both a predictive and financial viability standpoint.

5.4 Performance Comparison

The hybrid models significantly outperformed their standalone counterparts. For instance, the **FinBERT + TCN** configuration achieved an **MSE of 0.004**, **accuracy of 88%**, and **F1-score of 0.86**, outperforming LSTM-only and ARIMA-only models which lagged in both precision and recall. **Late fusion** slightly underperformed compared to **attention-based intermediate fusion**, indicating the importance of learning contextual weightings. Additionally, the hybrid models achieved a **Sharpe ratio improvement of 22%** over traditional advisory signals.

5.5 Case Studies

A case study involving **Tesla (TSLA)** during its Q3 earnings call in 2021 demonstrated the hybrid model's ability to anticipate a 12% price spike. The NLP module detected strong positive sentiment and earnings beat, while the time-series model captured an upward trend. The fusion model issued a **buy signal** a full day before the surge. Similar success was noted in identifying a downward trend in **Meta (META)** stock following negative sentiment post a privacy policy change.

6. DISCUSSION

6.1 Interpretability and Explainability

One of the core challenges of deep learning models in finance is the "black-box" nature of their decisions. To address this, we employed **SHAP (SHapley Additive exPlanations)** values and **attention weight visualizations** in both the NLP and time-series modules. This allowed us to highlight which words, dates, or indicators influenced the model's output. Such tools make the hybrid model explainable and more acceptable in regulatory and institutional environments.

6.2 Challenges

Despite promising results, several challenges persist:

- **Data alignment** between sentiment and price timelines is non-trivial due to news delays and market hours.
- **Model drift** in live financial markets can quickly degrade model performance, requiring constant retraining.
- **Overfitting risk** is high in financial datasets due to noise and low signal-to-noise ratio.
- **Text ambiguity** and sarcasm in social media require sophisticated NLP capabilities beyond sentiment polarity.

6.3 Potential Enhancements

To further improve the system:

- **Incorporating reinforcement learning** for real-time trading strategies.
- Use of **multimodal transformers** that jointly process text and time-series data.
- Adding **macroeconomic indicators**, such as GDP forecasts or interest rates, for broader context.
- Using **active learning** to continuously fine-tune sentiment models with new market data.

7. CONCLUSION AND FUTURE WORK

This paper proposed and validated a hybrid AI framework that combines **financial NLP** with **deep time-series forecasting models** for stock advisory applications. Experimental results show that hybrid models significantly outperform standalone models in both predictive accuracy and financial performance metrics. By integrating real-time sentiment signals with price trends, the system provides a more holistic and robust foundation for intelligent investment decision-making.

In the future, we aim to extend this framework to **other asset classes** (e.g., crypto, commodities), introduce **online learning capabilities**, and develop **personalized advisory agents** that can tailor recommendations to individual risk profiles. Additionally, exploring **federated learning** for decentralized financial environments offers an exciting avenue for secure and scalable deployment.

REFERENCES

[1] Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.

[2] Apu, K. U., Rahman, M. M., & Hoque, A. B. (2022). Forecasting future investment value with machine learning, neural networks, and ensemble learning: A meta-analytic study. *Review of Applied Science and Technology*, 4(2), 13–28. https://rast-journal.org

[3] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1–8.

[4] Box, G. E. P., & Jenkins, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). Wiley.

[5] Cao, L. (2022). AI in finance: Challenges, techniques, and opportunities. *ACM Computing Surveys, 55*(4), Article 88. https://doi.org/10.1145/3502289

[6] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383–417.

[7] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research, 270*(2), 654–669.

[8] Ko, C. R., & Chang, H. T. (2021). LSTM-based sentiment analysis for stock price forecast. *PeerJ Computer Science, 7*, e408. https://doi.org/10.7717/peerj-cs.408

[9] Leow, E. K. W., Nguyen, B. P., & Chua, M. C. H. (2021). Robo-advisor using genetic

algorithm and BERT sentiments from tweets for hybrid portfolio optimisation. *Expert Systems with Applications, 185*, 115652. https://doi.org/10.1016/j.eswa.2021.115652

[10] Lim, B., Arik, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting, 37*(4), 1748–1764.

[11] Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications, 41*(16), 7653–7670.

[12] Ni, H., Wang, S., & Cheng, P. (2021). A hybrid approach for stock trend prediction based on tweets embedding and historical prices. *World Wide Web, 24*(5), 1523–1541. https://doi.org/10.1007/s11280-021-00880-9

[13] Qiu, J., Wang, B., & Zhou, C. (2020). Forecasting stock prices with long-short term memory neural network based on attention mechanism. *PLOS ONE, 15*(1), e0227222. https://doi.org/10.1371/journal.pone.0227222

[14] Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance, 62*(3), 1139–1168.

[15] Sesen, M. B., Romahi, Y., & Li, V. (2019). Natural language processing of financial news. In T. Guida (Ed.), *Big Data and Machine Learning in Quantitative Investment* (pp. 185–202). Wiley.