RESEARCH ARTICLE OPEN ACCESS

A Critical Review of Modern Techniques for Stock Market Prediction

Pavankumar Patel*, Krina Masharu**
*B Tech in Computer Science Engineering, Atmiya University, Rajkot, India
Email: pavankumarpatel169@gmail.com
**Faculty of Engineering and Technology (CE), Atmiya University, Rajkot, India
Email: krina.masharu@atmiyauni.ac.in

--************************----------------------------------

Abstract:

The stock market has captivated academics for years because of its dynamic, nonlinear, and seemingly random behavior. Earlier research using statistical and econometric models has yielded to machine learning (ML), deep learning (DL), graph neural networks (GNNs), and large language models (LLMs) in the recent literature. In this review, I provide a reflection on more than fifteen of the last several years' papers from 2019-2025 that consider the emerging empirical approaches and the increasing inclusion of alternative data sources, such as sentiment data, in their studies, alongside addressing the underlying issues of reproducibility and explainability of how or why models produce their respective predictions. I observed that while the novel approaches generally report improved predictive performance over baseline models, these results are often based on unrealistic evaluation conditions. The primary takeaway from my reflection is that hybridized profiling approaches (econometric energy, ML exploratory flexibility, and alternative context data) seem to be the best suited for future research on stock market predictions, while the practical deployment of hybrid systems remains limited due to nonstationarity (union and/or intersection of spatial and temporal data/event characteristics) and a lack of transparency in the evaluations of the models.

Keywords: **Stock market prediction, econometrics, machine learning, deep learning, sentiment, graph neural networks, large language models**.

--************************----------------------------------

I. INTRODUCTION

The stock market is one of the most widely studied and least predictable systems. For decades, researchers have debated whether markets actually display efficiency (as outlined in the Efficient Market Hypothesis) and whether patterns truly exist that can be exploited. Upon reviewing the literature, it is clear that while absolute predictability is not yet here, methodological developments have led us in the direction of probabilistic signals being identified and used more skilfully.

Historically, researchers relied upon linear statistical models, as these were straight-forward, interpretable, and obedient to mathematical elegance, and did underestimate the risk posed to either volatile periods or periods of crises. Rapid advancements in the computation resources and data enabled the

discipline to move to broad-based machine learning in the 2010s and then to deep learning in the 2020s.

One major shift is the nature of the data being used. In the early years of the research, the majority of data was historical prices and volumes. Today, researchers are including textual sentiment data (news, tweets, reports), correlation networks, and even unorthodox data, such as Google searches and satellite imaging. I see this as an acknowledgement that the markets are not just numerical—they are socio-economic systems captured by human psychology. This paper reviewed these in-depth. I certainly do not present studies in a linear, mechanical way, but critically reflect on their merits, limitations, and relative positioning.

II. BACKGROUND AND TRADITIONAL APPROACHES

As I started reflecting on prior approaches, I realized that ARIMA, GARCH, and their offshoots have been the backbone of stock forecasting for several decades. These methods assume linearity and stationarity, which makes estimation easier, yet they are optimistic when applied to real markets. For those familiar with market models, GARCH can account for volatility clustering but fails to adequately capture shifts in regimes (for example, the 2008 financial crisis or the COVID-19 pandemic).

Surprisingly, recent articles [6-7] continue to apply ARIMA/GARCH as references when testing new models. While these models do not always represent the top-performing approach, they do provide a "baseline" in terms of their interpretations or stability. Several studies are even hybridized to incorporate ARIMA while utilizing ML/DL criteria for their other predictions, where the authors interpret ARIMA as being the linear portion of the forecasting and the ML models representing the non-linear residuals.

Overall, I seem to take away that traditional models have not become obsolete, simply incomplete. There is still value in using traditional models for forecasting, especially with models meant to explain variations in volatility and for interpretations, but they can't solely exist in the realm of forecasting.

III. MACHINE LEARNING APPROACHES

The emergence of machine learning (ML) infused the enterprise with enthusiasm. Algorithms such as Support Vector Machines (SVM), Random Forest, and gradient boosting (e.g., XGBoost and LightGBM) quickly garnered attention. In functional reviews of ML-centered papers [8–9], a few points stuck out:

- **Non-linearity handling:** ML can capture non-linear relationships, which ARIMA cannot.
- **Ensemble superiority:** Almost all articles reported superior results in experiments including ensembles (e.g., XGBoost) than using single models.
- **Feature engineering as a bottleneck:** ML models require features to be designed for input into them—moving averages, momentum indicators, macroeconomic features.
- **Overestimated performance:** Years later, many papers exhibited 90% or higher accuracy rates, but were later caught using random train-test split methods, allowing future information to "leak" into the training task. A forthcoming review (Springer 2025) notes that this should be posted against, noting that the true predictive power may be lower than shown.

I left this area feeling mixed about ML. On one hand, this method (ML) represented improvements compared to relevant "traditional models". On the other hand, its dependence on human features and other evaluation mistakes diminishes the faith we could have in those evaluations.

IV. DEEP LEARNING AND HYBRID MODELS

Deep learning changed what was possible dramatically. Recurrent Neural Networks (RNNs), and in particular Long Short-Term Memories (LSTMs) and Gated Recurrent Units (GRUs), are best equipped to deal with time series data. A number of papers [10-11] noted that LSTMs were more likely to have correctly predicted the short-term direction of price changes than SVMs or ARIMA. Convolutional Neural Networks (CNNs) have also been used by more closely dealing with stock sequences as image-like structures.

While the benefits are evident, DL is far from a silver bullet. My observations:

- **Risk of overfitting:** LSTMs memorize noise unless they are explicitly regularized.

- **Confusion of the market regime:** Models learned during a bull market do not easily transfer to a crisis.
- **Sensitivity to hyperparameters:** The results dramatically changed with even minimal modifications to learning rate or window size.

The most thrilling thing I learned about were hybrid models. For instance, a paper (ScienceDirect, 2025) coupled GARCH (volatility) with LSTM (nonlinear dependencies). Another paper coupled CNNs with sentiment analysis and demonstrated higher accuracy on short-term price forecasts. These hybrid forms are what I see in the future of our field: integration rather than isolation. Rather than discarding or depending too heavily on a single "magic model", researchers are blending the complementary and distinct strengths of various models. The only negative is greater complexity and less explainability.

V. THE ROLE OF SENTIMENT, TEXT, AND ALTERNATIVE DATA

If there is anything that kept recurring and caught my interest, that is the significance of sentiment power. The markets run not only on numbers but on expectations and sentiment. Surveys conducted lately [12–14] confirmed how the inclusion of news headlines, accounting reports, and tweets enhances the forecast.
Some of my review's important points:

- **Short-term benefit:** Extremes benefit most during day or intraday forecasting, when sentiment shifts significantly.
- **Data alignment problem:** Most research aligns news dates with wrong price reactions or doubts results. The ones using high-frequency alignment were much more convincing.
- **Noisy sources:** Twitter and Reddit contribute with signals, but only through filtering out spam and untrustworthy accounts.

In addition to text, scholars have dabbled with extraneous data: Google Trends search volumes, even satellite photographs to guesstimate retail activity. Such data is expensive and not open to replication very often, and that's a deterrent to me. My general feeling: sentiment and alternative data provide models with greater realism, but with caution. Bad preprocessing can contribute to more noise than signal.

VI. NEW DIRECTIONS: GRAPH NEURAL NETWORKS AND LARGE LANGUAGE MODELS

The recent years (2023–2025) have seen a surge in graph-based approaches. Stocks don't stand alone—they are connected by industry, sector, supply chain, and correlation. Graph Neural Networks (GNNs) utilize this topology best. GNNs have proven to beat LSTMs in most applications of cross-sectional prediction (simultaneous prediction of multiple stocks), according to ACM and ScienceDirect polls [15].

Another potential opportunity is using Large Language Models (LLMs). Some experimentation [16] with GPT-type transformers for financial reporting, not only extracting sentiment but structural knowledge (e.g., risk mentions in earnings calls), has shown promise. LLMs are not without their problems, however:

- They will occasionally "hallucinate".
- They possess no foundation of financial understanding.
- They can be over-sensitive to the wording of prompts.
- In my view, LLMs are better understood as feature generators (translating unstructured text to actionable signals) rather than predictors.

VII. CHALLENGES AND OPEN ISSUES

Despite progress, several objections exist:

- **Non-stationarity:** Markets evolve through time; models educated on historical data tend to perform badly in new environments.
- **Overfitting and testing bias:** Much prior work overestimates accuracy by using random splits or omitting transaction costs.

- **Interpretability:** DL, GNNs, and LLMs are not explainable—an ill portent for finance.
- **Reproducibility:** Secret data and lack of open code preclude verification.
- **Practicality gap:** Theoretical models do not account for slippage, commissions, and liquidity—theory models untradeable models.

In my opinion, the solution to these involves having benchmark datasets, standardization of backtesting protocols, and adopting explainable AI for finance.

VIII. CONCLUSION

Stock market modeling has evolved over time from simple but elegant econometric models to computationally heavy but successful AI-driven approaches. My scan of 15+ current articles suggests a pair of self-evident trends:

- No single model reigns supreme across all environments.
- Hybrid models that blend econometrics, ML, DL, and sentiment work best.
- Alternative data (news, texts, social media) generate value but must be properly matched.

The biggest challenges are not technical correctness, but rigor of evaluation, interpretability, and extrinsic applicability. In looking to this field, I'm comforted that the future is integrative: combining quantitative, textual, and structural data within rigorous assessment frameworks. Stock market prediction may never be perfect—but with judicious research, it can become more reliable, transparent, and informative to aid decisions.

- **Ensemble superiority:** Almost all articles reported superior results in experiments including ensembles (e.g., XGBoost) than using single models.
- **Feature engineering as a bottleneck:** ML models require features to be designed for input into them—moving averages, momentum indicators, macroeconomic features.
- **Overestimated performance:** Years later, many papers exhibited 90% or higher accuracy rates, but were later caught using random train-test split methods, allowing future information to “leak” into the training task. A forthcoming review (Springer 2025) notes that this should be posted against, noting that the true predictive power may be lower than shown.

I left this area feeling mixed about ML. On one hand, this method (ML) represented improvements compared to relevant "traditional models". On the other hand, its dependence on human features and other evaluation mistakes diminishes the faith we could have in those evaluations.

References:

1. **SAGE (2024).** *Bibliometric literature review of stock price forecasting.*
2. **ScienceDirect (2024).** *Stock Market Prediction Using Artificial Intelligence: A Systematic Review.*
3. **MDPI (2023).** *Forecasting Stock Market Prices Using Machine Learning.*
4. **ScienceDirect (2022).** *Systematic Review of Stock Market Prediction.*
5. **arXiv (2024).** *Study on Stock Forecasting Using Deep Learning.*
6. **ACM (2024).** *Graph Neural Networks for Financial Prediction: A Survey.*
7. **ScienceDirect (2025).** *Key Technical Indicators for Stock Prediction.*
8. **AIP (2024).** *Comprehensive Survey of Stock Market Prediction Methods.*
9. **MDPI (2023).** *Machine Learning Approaches in Stock Market Prediction.*
10. **ScienceDirect (2024).** *Role of Sentiment in Stock Market Forecasting.*

11. **arXiv (2024).** *Deep Learning in Stock Market: A Surve.*
12. **JISAE (2024).** *Systematic Review on DL-based Stock Prediction with News Sentiment.*
13. **ScienceDirect (2025).** *Data-driven Models for Stock Forecasting.*
14. **arXiv (2024).** *Evaluation of ML/DL Methods for Stock Market Prediction.*