Evaluation of Stock Market Prediction Techniques

Mr. Abhishek Malaviya
School of Computing
Dublin City University
Dublin, Ireland
abhishek.malaviya2@mail.dcu.ie

Dr. Andrew McCarren
School of Computing
Dublin City University
Dublin, Ireland
Head of School of Computing
andrew.mccarren@dcu.ie

I. Introduction

Financial time series data, such as stock prices, are inherently volatile, characterized by noise, non-stationarity, and sensitivity to external factors [1], [2] like economic events and investor sentiment. These challenges render traditional prediction methods such as linear regression or ARIMA [3], [4] are ineffective in capturing dynamic market behavior. The research question driving this review is:

How can we predict highly volatile signals, such as stock market movements, amidst noisy and chaotic data?

To address this, we conducted a systematic search using queries including:

- *Systematic Literature Review on volatility*
- *Systematic Literature Review on ARIMA stock market*
- *Systematic Literature Review on Volatility and signal detection*
- *Systematic Literature Review on Stock Market Prediction*
- *Systematic Literature Review on Temporal Convolutional Networks and Forecasting*
- *Stock Price Prediction using Neural Networks*

This review synthesizes findings from 20 studies, to evaluate methodologies for modeling and predicting volatile financial data.

II. Historical Methods

Historical statistical methods remain foundational for short-term forecasting:

ARIMA combines autoregressive and moving average components to model trends and seasonality. While effective in stable conditions (e.g., achieving 97.6% accuracy [3]), it struggles with sudden market shifts and external factors. Exponential Smoothing prioritizes recent observations, adapting to gradual volatility changes but lacking robustness in turbulent markets. Linear Regression assumes linear relationships between variables, often failing to capture the non-linear dynamics of financial data [4].

These methods are computationally efficient but limited by their inability to incorporate qualitative data or adapt to rapid volatility spikes.

Recent studies, such as [5], leverage wavelet transforms to decompose non-stationary data into trend and seasonal components, mitigating noise and improving forecasting robustness.

III. Measuring Volatility

Volatility modeling is critical for risk management and prediction accuracy. Key approaches include:

- GARCH-family models [1]: Explicitly model volatility clustering and asymmetry. EGARCH and GJR-GARCH address *leverage effects*, where negative shocks amplify volatility. These models excel in capturing conditional variance but require parameter tuning for different market regimes.
- Time Series Decomposition: ARIMA implicitly addresses volatility through residual analysis [3], while wavelet transformations denoise data to isolate trends [6].
- Machine Learning: LSTM networks indirectly model volatility via temporal dependencies [7], [8], while SVM-based methods classify trends without explicit volatility modeling [9].
- Hybrid Frameworks: [2] highlights neural networks (MLPs, RNNs) combined with GARCH to enhance volatility forecasting, though standardization remains a challenge.

The new systematic reviews [1], [2] and [4] emphasize that while GARCH models dominate explicit volatility analysis, hybrid architectures integrating AI and econometrics show promise for future research.

Wavelet decomposition (e.g., MRA-WT) isolates high-frequency noise and low-frequency trends in volatile data, as demonstrated in vegetation forecasting [5]. This approach could enhance financial volatility modeling by separating chaotic fluctuations from underlying trends.

IV. Literature Review

Some of the latest techniques in time series prediction have leveraged deep learning architectures. For example, one study [10] uses a Transformer-based deep reinforcement learning agent combined with the Black-Litterman model. This method trains on data from the Dow Jones Industrial Average to capture dynamic asset correlations while incorporating investor views to adjust predictions. The result is a strategy that outperforms traditional methods by generating 42% higher returns.

Another approach [6] introduces a hybrid model that combines 3D convolutional neural networks (3D-CNN) with gated recurrent units (GRU). In this model, the 3D-CNN is used to extract spatial features from financial data, while the GRU captures temporal dependencies. To address the high noise levels typically found in financial series, this method employs a wavelet transformation, which helps isolate important patterns by filtering out high-frequency noise. The model also incorporates a Dandelion Optimization Algorithm and a Blood Coagulation Algorithm for feature selection, achieving an impressive 99.14% accuracy on the NIFTY 50 dataset. However, its focus on numerical data means it may not fully account for qualitative influences.

Deep learning methods continue to evolve with models such as Bidirectional LSTM [7] that process data in both directions to capture past and future trends. By integrating technical indicators like the MACD and Money Flow Index, this approach achieves a high test accuracy of 96%, clearly outperforming simpler LSTM implementations. Similarly, an LSTM-based method [8] specifically targets the prediction of stock returns in the Chinese market, showcasing the LSTM's strength in capturing long-term dependencies despite the inherent noise. A different strategy [11] uses a standard Recurrent

Neural Network to model short-term price movements. While it is effective in the short term, it suffers from the common issue of vanishing gradients, which limits its ability to capture long-term patterns. Moreover, a study [12] that explores a combination of CNNs, LSTMs, and GANs demonstrates that integrating various deep architectures can reduce noise and capture complex trends, though it also faces challenges like overfitting. In other work [13], LSTM networks are used to forecast stock market indices with high accuracy, but the method remains sensitive to sudden market changes.

Another approach gaining traction in deep learning is [14] Residual Networks (ResNets), which address issues like the vanishing gradient problem in very deep networks. By using residual connections, ResNets allow for deeper architectures that can capture even more complex and abstract patterns in time series data, potentially improving predictions in volatile stock market conditions.

To further improve performance, several researchers have developed hybrid models that integrate deep learning with sentiment analysis. One hybrid approach [15] combines LSTM layers with sentiment scores derived from BERT, which processes financial news and social media posts. This method not only considers historical numerical data but also incorporates real-time textual information, leading to improved prediction accuracy—albeit with some variability across different stocks. Another study [16] takes a similar route by merging LSTM-based analysis with sentiment data from news articles, and it further enhances its predictions through ensemble learning techniques such as CatBoost and Random Forest, achieving very low error rates (MAE of 0.018 and RMSE of 0.015).

Additionally, an ensemble model [17] uses sentiment analysis to assess the polarity of news and combines this information with traditional ensemble learning models like Random Forest and Gradient Boosting.

A web crawler gathers real-time sentiment data which is then merged with historical price data to improve accuracy and robustness during volatile periods. Furthermore, a deep fusion model [18] integrates sentiment extracted from news headlines with LSTM-based time series analysis, demonstrating that a combination of qualitative and quantitative data can outperform many state-of-the-art models in trading simulations.

Not all researchers have relied on the complex architectures of deep learning. Traditional statistical methods remain relevant, particularly for short-term forecasts. One study [3] employs classic techniques such as ARIMA and Exponential Smoothing, prioritizing recent observations to capture short-term trends and achieving up to 97.6% accuracy in stable market conditions. Similarly, an RBF-SVM model [9] uses a Radial Basis Function kernel to map non-linear data into a higher-dimensional space, thus classifying stock price movements into upward and downward trends. Although this approach yields moderate accuracy (around 65.64%), it is mainly effective for short-term predictions.

Emerging architectures like Temporal Convolutional Networks (TCNs) combine dilated convolutions and residual blocks to capture long-term dependencies efficiently. Veg-W2TCN [5] outperformed LSTM and GRU in non-stationary forecasting, suggesting TCNs' potential for financial time series. Another such approach [19] applies soft computing techniques by integrating genetic algorithms, fuzzy reasoning, and semantic text mining to handle both quantitative financial indicators and qualitative news sentiment. This method takes advantage of soft computing's global search capabilities to better manage the stochastic nature of financial markets, even under high volatility, though it may face scalability issues. In another innovative study [20], federated learning is used to enhance an MLP–LSTM model. In this framework, local models are trained on proprietary data, and only model parameters are shared centrally. This approach not only achieves high accuracy (up to 98.3% with an RMSE as low as 0.0108) but also addresses the critical issue of data privacy.

V. Evaluation Techniques

Performance is assessed through:

- Accuracy Metrics: Hybrid models like 3D-CNN-GRU achieve up to 99.14% accuracy [6], while sentiment-LSTM ensembles report MAE as low as 0.018 [16].
- Risk-Adjusted Metrics: Studies incorporating portfolio optimization [10], [1] use Sharpe and Sortino ratios to evaluate volatility resilience.
- Error Rates: MAPE (less than 4% in Bi-LSTM [7]) and RMSE (0.0108 in federated learning [20]) quantify prediction stability.
- Scalability and Privacy: Federated learning and MLP-LSTM [20] address data privacy while maintaining accuracy.

[2] identifies a lack of standardized benchmarks, urging shared tasks (e.g., S&P 500 volatility forecasting) for equitable comparisons.

When comparing these methodologies, several points become clear. Deep learning models that rely solely on numerical data, such as the 3D-CNN-GRU [6], can achieve very high accuracy but might not adapt well to rapid changes caused by external factors. Hybrid models integrating wavelet decomposition with deep learning (e.g., Veg-W2TCN [5]) reduce computational complexity while improving accuracy, achieving 83% of pixels with RMSE less than 0.1 compared to 74% for standalone TCN. Other Hybrid models that combine numerical data with sentiment analysis [15], [16], [17], [18] tend to provide a more robust solution in volatile markets. Traditional statistical methods [3], [9] are effective for short-term forecasts but often lack the ability to incorporate real-time external information. Meanwhile, emerging paradigms like soft computing [19] and federated learning [20] show significant promise in addressing issues such as noise, scalability, and data privacy while maintaining high predictive performance. Wavelet-TCN hybrids achieve RMSE reductions of 5–20% over standalone models [5], demonstrating the value of decomposition in noisy environments. Similar approaches could stabilize financial predictions under volatility.

- Best Performers: Hybrid models (e.g., sentiment-LSTM [15], 3D-CNN-GRU [6]) and transformer-DRL combinations [10] excel in volatile markets.
- Latest Trends: Federated learning [20] and transformer-based frameworks [10], [4] address scalability and modern market dynamics.
- Persistent Challenges: Overfitting, computational complexity, and data heterogeneity remain hurdles.

Overall, the literature reveals that the most effective stock prediction methods are those that balance numerical analysis with the integration of qualitative data, such as news sentiment. While traditional methods remain useful, the latest techniques in deep learning, hybrid models, and emerging paradigms provide improved accuracy and adaptability to the noisy, volatile nature of financial markets. Evaluation techniques in these studies typically include accuracy percentages, error metrics like RMSE and MAE, and assessments of scalability and data privacy. Future research should continue to explore multimodal data integration, enhanced privacy-preserving techniques, and causal reasoning to further improve prediction robustness and real-world applicability. Incorporating wavelet decomposition and TCN architectures—as seen in Veg-W2TCN [5]—could address computational bottlenecks while improving robustness to non-stationarity in financial data

Acknowledgments

The author utilized ChatGPT (OpenAI) in assisting with grammatical corrections and paraphrasing when compiling this literature review. Any final content was personally reviewed and edited by author to ensure accuracy, coherence, and alignment with the research objectives [21].

References

[1] R. Bhowmik and S. Wang, "Stock market volatility and return analysis: A systematic literature review," *Entropy*, vol. 22, no. 5, 2020, h-Index: 101. [Online]. Available: https://www.mdpi.com/1099-4300/22/5/522

[2] W. Ge, P. Lalbakhsh, L. Isai, A. Lenskiy, and H. Suominen, "Neural network–based financial volatility forecasting: A systematic review," *ACM Computing Survey*, vol. 55, no. 1, Jan. 2022, 213. [Online]. Available: https://doi.org/10.1145/3483596

[3] M. P.Ramu, M. A.Vani Priya, M. D.Roopa Sree, and M. S.Sankeerthana, "Stock price prediction using time series," *International Research Journal of Engineering and Technology (IRJET)*, vol. 09, no. 06, 2022, impact Factor: 8.226. [Online]. Available: https://www.irjet.net/archives/V9/i6/IRJET-V9I6210.pdf

[4] D. P. Gandhmal and K. Kumar, "Systematic analysis and review of stock market prediction techniques," *Computer Science Review*, vol. 34, p. 100190, 2019, h-Index: 75. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S157401371930084X

[5] M. Rhif, A. B. Abbes, B. Martínez, and I. R. Farah, "Veg-w2tcn: A parallel hybrid forecasting framework for non-stationary time series using wavelet and temporal convolution network model," *Applied Soft Computing*, vol. 137, p. 110172, 2023, h-Index: 190. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494623001904

[6] B. Jagadesh, N. RajaSekhar Reddy, P. Udayaraju *et al.*, "Enhanced stock market forecasting using dandelion optimization-driven 3d-cnn-gru classification," *Scientific Reports*, vol. 14, p. 20908, 2024, h-Index: 315. [Online]. Available: https://doi.org/10.1038/s41598-024-71873-7

[7] J. Shah, R. Jain, V. Jolly, and A. Godbole, "Stock market prediction using bi-directional lstm," in *2021 International Conference on Communication information and Computing Technology (ICCICT)*, 2021, pp. 1–5, h-Index: N/A. [Online]. Available: https://10.1109/ICCICT50803.2021.9510147

[8] K. Chen, Y. Zhou, and F. Dai, "A lstm-based method for stock returns prediction: A case study of china stock market," in *2015 IEEE International Conference on Big Data (Big Data)*, 2015, pp. 2823–2824, h-Index: 31. [Online]. Available: https://10.1109/BigData.2015.7364089

[9] Z. Liu, Z. Dang, and J. Yu, "Stock price prediction model based on rbf-svm algorithm," in *2020 International Conference on Computer Engineering and Intelligent Control (ICCEIC)*, 2020, pp. 124–127, h-Index: N/A. [Online]. Available: https://doi.org/10.1109/ICCEIC51584.2020.00032

[10] R. Sun, A. Stefanidis, Z. Jiang, and J. Su, "Combining transformer based deep reinforcement learning with black-litterman model for portfolio optimization," *Neural Comput. Appl.*, vol. 36, no. 32, p. 20111–20146, Aug. 2024, h-Index: 130. [Online]. Available: https://doi.org/10.1007/s00521-024-09805-9

[11] K. R. Reddy, B. T. Kumar, V. R. Ganesh, P. Swetha, and P. K. Sarangi, "Stock market prediction using recurrent neural network," in *2022 IEEE International Conference on Current Development in Engineering and Technology (CCET)*, 2022, pp. 1–4, h-Index: N/A. [Online]. Available: https://10.1109/CCET56606.2022.10080154

[12] G. Duan, M. Lin, H. Wang, and Z. Xu, "Deep neural networks for stock price prediction," in *2022 14th International Conference on Computer Research and Development (ICCRD)*, 2022, pp. 65–68, h-Index: N/A. [Online]. Available: https://doi.org/10.1109/ICCRD54409.2022.9730340

[13] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, K. R. Dahal, and R. K. Khatri, "Predicting stock market index using lstm," *Machine Learning with Applications*, vol. 9, p. 100320, 2022, h-Index: N/A. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666827022000378

[14] Y. Jia, A. Anaissi, and B. Suleiman, "Resnls: An improved model for stock price forecasting," *Computational Intelligence*, vol. 40, no. 1, p. e12608, 2024, h-Index: N/A. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/coin.12608

[15] X. Weng, X. Lin, and S. Zhao, "Stock price prediction based on lstm and bert," in *2022 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2022, pp. 12–17, h-Index: N/A. [Online]. Available: https://10.1109/ICMLC56445.2022.9941293

[16] A. Churi, D. Chakraborty, R. Khatwani, G. Pinto, P. Shah, and R. Sekhar, "Stock price prediction using deep learning and sentiment analysis," in *2023 2nd International Conference on Futuristic Technologies (INCOFT)*, 2023, pp. 1–6, h-Index: N/A. [Online]. Available: https://doi.org/10.1109/INCOFT60753.2023.10425124

[17] S. K. Bharti, P. Tratiya, and R. K. Gupta, "Stock market price prediction through news sentiment analysis & ensemble learning," in *2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC)*, 2022, pp. 1–5, h-Index: N/A. [Online]. Available: https://doi.org/10.1109/iSSSC56467.2022.10051623

[18] P. Chen, Z. Boukouvalas, and R. Corizzo, "A deep fusion model for stock market prediction with news headlines and time series data," *Neural Computing and Applications*, vol. 36, pp. 21 229–21 271, 2024, h-Index: 130. [Online]. Available: https://doi.org/10.1007/s00521-024-10303-1

[19] J. Padilla-Caballero, R. Diaz-Quichiz, M. Quispe-Huayllapuma, H. Huayllapuma-Rivera, L. Rojas-Zuñiga, and F. Cardenas-Palomino, "Recent developments and methodologies for stock market prediction using soft computing technique," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, vol. 6, 2023, pp. 1781–1785, h-Index: N/A.

[20] J. Kumarappan, E. Rajasekar, S. Vairavasundaram *et al.*, "Federated learning enhanced mlp–lstm modeling in an integrated deep learning pipeline for stock market prediction," *International Journal of Computational Intelligence Systems*, vol. 17, p. 267, 2024, h-Index: 56. [Online]. Available: https://doi.org/10.1007/s44196-024-00680-9

[21] OpenAI, "Chatgpt," https://chat.openai.com, 2023, accessed: 20th February 2025. Used for summarization and content refinement.