International Journal of Applied and Advanced Multidisciplinary Research (IJAAMR)
Vol. 1, No. 2, 2023: 215 - 222

Analytical Review of Machine Learning Algorithms (Models) for Stock Market Prediction

Dr. M.A. Thalor[1]*, Tejas Bose[2],
AISSMS IOIT Pune India.
Corresponding Author: Tejas Bose tejasrbose@gmail.com

A R T I C L E I N F O

Keywords: Stock Market Forecasting, Asset Price Prediction, Computational Intelligence, Deep Learning Models, Artificial Neural Networks (ANNs)

Received : 06, August
Revised : 16, September
Accepted: 26, October

A B S T R A C T

The realm of stock market forecasting presents a formidable challenge, given the intricate, noisy, chaotic, and ever-evolving nature of its time series data. However, the advent of computational advancements offers a ray of hope, as intelligent models hold the potential to assist investors and analysts in mitigating the inherent risks associated with financial markets. In recent years, Deep Learning models have garnered significant attention, with numerous studies delving into their application for predicting stock prices using historical data and technical indicators Yet, the ultimate goal in this pursuit is not merely prediction but validation, a crucial step in the context of the financial markets. With a significant 73.5% of the research in this field using the LSTM (Long Short-Term Memory) technique, it is clear from this systematic review's conclusion that this technique is the best. The review does, however, point out significant gaps in the literature, with just 35.3% of studies covering profitability measurements and only two publications going into great detail into risk management

DOI: https://doi.org/10.59890/ijaamr.v1i2.517

INTRODUCTION

Predicting asset prices in the stock market remains an intricate and challenging endeavour, influenced by a multitude of micro and macroeconomic factors, including political events, news, and financial data. These complex dynamics render the task inherently nonlinear and non stationary, intensifying the complexities involved. To tackle this complexity, researchers employ two primary approaches: Fundamental Analysis (FA) and Technical Analysis (TA). FA delves into a company's data to assess its long-term growth potential, while TA relies on historical price and volume data, ignoring company-specific information. Technical analysts harness various tools such as Technical Indicators (TIs) and candlestick pattern analysis to forecast price movements. Recent advancements in computational intelligence, particularly Deep Learning (DL), have gained prominence in stock market forecasting. DL models, such as Artificial Neural Networks (ANNs), have demonstrated superior performance compared to traditional statistical methods. However, the challenge lies in identifying the most effective set of TIs for these models. This research paper builds upon this foundation by focusing on DL techniques for stock market price forecasting. It seeks to evaluate the accuracy and profitability metrics employed to validate these models and explores the trading strategies adopted. In an ever- evolving financial landscape, understanding the potential of DL and its impact on stock market prediction is crucial for informed decision-making and investment strategies.

LITERATURE REVIEW

We have analyzed various articles based on stock market prediction, in that every article has different approach for stock price prediction but in common algorithms used in their implementation are some how some frequently using algorithms. Based on different factors we analyzed the articles which are based on stock market forecasting. Basically we performed comparative study of the various articles from this we get that, below study In a comparative study conducted by [1], the inclusion of Technical Indicators (TIs) was found to significantly enhance the predictive power of machine learning (ML) models for financial forecasting. TIs were derived from 1-hour intraday data over a 24-hour forecast window.

Data preprocessing involved normalization and the use of an autocorrelation function to select pertinent input data, resulting in the creation of 9 TIs. The study implemented 14 ML models, including Convolutional Neural Networks (CNNs), and the results demonstrated a notable improvement in next-day price forecast accuracy attributed to the incorporation of Tis Innovatively, [2] introduced a novel model integrating Convolutional Neural Networks (CNNs) with graph theory principles. This approach was tested through two distinct models— Correlation is the basis for one, and causation for the other. In addition, the study used evaluation metrics including mean absolute percentage error (MAPE), mean mean-square error (RMSE), and mean absolute error (MAE) to compare these models with conventional methods like linear regression and ARIMA.

The results showed that, notably, the suggested CNN-based model performed better in terms of prediction accuracy than traditional approaches. Notably, neither the performance of an LSTM network nor a basic CNN network was examined in the study, underscoring the unique contribution of this novel CNN-graph theory hybrid. While a majority of the analyzed publications employed Long Short-Term Memory (LSTM) techniques for financial forecasting, several commonalities emerged in terms of pre processing, result comparisons, and accuracy metrics. Notably, authors such as [3], [4], [5], [6], and [7] utilized asset prices and Technical Indicators (TIs) as network attributes. They followed a normalization process to prepare the data for input into LSTM networks. However, it is worth highlighting the work of Qiu et al. [7], who introduced an innovative approach by combining LSTM with Gated Recurrent Unit (GRU). This hybrid model aimed to harness LSTM's expertise in processing sequential data while leveraging the simplicity of GRU, resulting in reduced training time and computational cost. This novel approach presents exciting possibilities for more efficient financial forecasting using recurrent neural networks

Innovative LSTM-based approaches for financial forecasting have shown promise. Auto SLSTM [8] incorporated an autoencoder into the LSTM network, reducing input data noise and outperforming traditional LSTM and MLP models. It was observed that higher parameter values led to increased error accumulation in varying forecast horizons. Another unique model [9] stacked two LSTM layers, generating 400 market based characteristics. Notably, it employed dynamic attribute selection during training, resulting in varied training data. Additionally, [10] optimized LSTM by using the correlation tensor technique to create adaptive Technical Indicators (TIs), improving accuracy and reducing MSE. These novel strategies hold potential for enhancing LSTM-based financial forecasting Incorporating textual data, such as news articles, into financial forecasting models has shown promising results in recent studies [11], [12], [13], [14]. These models leveraged LSTM, autoencoders, deep belief networks (DBN), and attention mechanisms (AM) to process textual information. A crucial step in these models was the pre-processing of textual data using sentiment analysis techniques. This processed textual data was then integrated with price and Technical Indicator (TI) data. Moreover, researchers have explored hybrid forecasting models that combine LSTM or RNN layers with CNN layers to extract valuable information from textual sources [15], [16], [17], [18], [19]. By extracting patterns from news channels and social networks, these models generated richer information beyond asset prices alone. Remarkably, all of these hybrid models outperformed models relying solely on LSTM implementations. These findings underscore the potential of incorporating textual data and hybrid models to enhance financial forecasting accuracy and provide deeper insights into market dynamics Financial forecasting models employ various trading strategies. Traditional rules trigger buys and sells based on model signals. Some include a 'neutral' class. Longer-term strategies hold positions for days, while Day Trading (DT) strategies operate on shorter timeframes. Unique approaches focus on asset appreciation and devaluation predictions for trading decisions, offering diverse strategies for financial forecasting. In financial forecasting for the stock market, measuring profitability is critical. Accuracy and precision do not necessarily guarantee profitability. Many studies focus on gross profit, neglecting operating costs and fees, leading to misleading conclusions. Some consider ROI (return on investment) or net profitability after accounting for costs, revealing the real financial performance. It's essential to consider costs when evaluating forecasting models to make informed decisions. Wang et al. [21] introduced the Weighted F- Score metric, tailored for financial markets, acknowledging that different types of forecasting errors impact financial performance differently. Additionally, Wang et al. accounted for slippage, a crucial factor in trading performance often overlooked in other studies Risk management in stock market forecasting involves the implementation of techniques to mitigate capital loss (stop loss, SL) or maximize profit (take profit, TP). However, it's noteworthy that only two publications in our study addressed these aspects. In the study by [12], two distinct thresholds for TP and SL were employed: 1% and 2%. Trading strategies executed long or short positions at the trading session opening, concluding the transaction by the end of the same day. However, if the asset's price movement exceeded the defined threshold of 1% or 2%, either TP or SL was triggered accordingly. On the other hand, Song and Lee [22] employed an SL of 12% and TP of 24%, coupled with a maximum days positioned parameter. For example, a maximum of ten days positioned was set, and if the asset didn't reach the SL or TP within this period, the operation was closed.

METHODOLOGY

Hybrid Deep Learning Models

Consider developing a novel hybrid deep learning model that combines the strengths of different deep learning architectures, such as combining Convolutional Neural Networks (CNNs) with Recurrent Neural Networks (RNNs) like Long Short-Term Memory (LSTM). Evaluate how this hybrid model performs compared to single-model approaches.

Advanced Technical Indicators

Explore the creation of new and advanced technical indicators that can capture subtle market trends and patterns. These indicators can be designed specifically to leverage the strengths of deep learning models. Evaluate the effectiveness of these custom indicators in improving forecasting accuracy.

Transfer Learning

Investigate the application of transfer learning techniques in stock market forecasting. Pre-trained deep learning models, such as those trained on vast financial datasets, can be fine tuned for specific forecasting tasks. Assess whether transfer learning can improve prediction performance.

Sentiment Analysis

Incorporate sentiment analysis of financial news and social media data into your model. Assess how sentiment analysis can enhance stock market predictions when combined with technical analysis and deep learning.

Textual Data Integration

Several studies incorporated textual data, such as news and sentiment analysis, into LSTM-based models. These models concatenated textual data with price and TI data, achieving favorable result

Hybrid Models

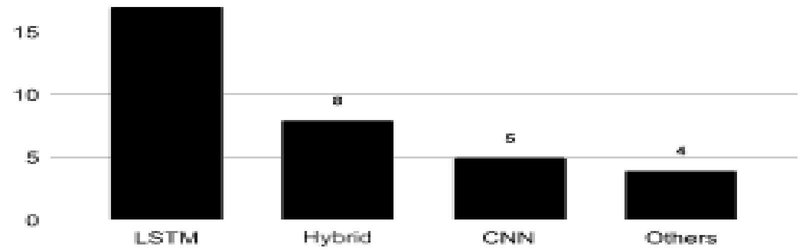
Hybrid models combining LSTM or RNN layers with CNN layers were explored. These models extracted patterns from textual data, such as news channels and social networks, resulting in improved predictions compared to models with LSTM alone.

Optimization Algorithms

Some studies introduced optimization algorithms like the Rider-based monarch butterfly optimization and GAN (Generative Adversarial Network) techniques to improve training and forecasting processes.

RESEARCH RESULT & DISCUSSION

The LSTM network is the preferred choice for time series forecasting in stock market analysis due to its memory storage capabilities and ability to address the vanishing gradient problem. A majority of studies, 73.5% to be precise, rely on LSTM-based models, making it the dominant deep learning technique in this domain. Additionally, Python and TensorFlow are the go-to tools for developing predictors using historical price data. These tools are complemented by other popular libraries like NumPy, Pandas, Scikit-Learn, Keras, TA-Lib, and TA4J for generating technical indicators. Researchers in this field commonly leverage this tech stack for robust and effective forecasting models. This information highlights the prevalence of LSTM networks and the standard tools used in stock market forecasting research.



Stock market forecasting research exhibits a global reach, encompassing various markets like North America, India, China, and cryptocurrencies. Yahoo Finance is the favored choice for historical stock price data due to its ease of use with Python libraries. For hybrid algorithms involving sentiment analysis and news data, researchers tap into sources such as Reuters, Bloomberg, and Twitter. Most studies employ daily data, while Day Trading (DT) is an alternative for those seeking ample training data and rigorous performance analysis. DT is especially suited for high-volatility assets, minimizing capital exposure and capitalizing on rapid market shifts."

CONCLUSION

In this comprehensive review of financial time series forecasting, we have analyzed 22 academic articles focusing on the intersection of deep learning (DL) techniques and technical analysis. Our analysis centered on four crucial aspects: predictor techniques, trading strategies, profitability metrics, and risk management. One prominent observation is the widespread adoption of the Long Short-Term Memory (LSTM) recurrent neural network due to its exceptional memory storage capabilities and effectiveness in addressing the vanishing gradient problem. Moreover, several hybrid models have emerged, leveraging LSTM for handling technical indicators and combining it with other techniques to incorporate news data, resulting in more robust forecasting outcomes. As we move forward, it is evident that there exist notable gaps in the existing literature, offering exciting opportunities for future research endeavors. These gaps encompass the development of hybrid models that seamlesslyintegrate both qualitative and quantitative input data, the exploration of intelligent and adaptive trading strategies, the formulation of metrics that establish a positive correlation between performance and profitability, and the implementation of robust risk management techniques. Addressing these research gaps will undoubtedly contribute to the continued evolution of financial time series forecasting.

REFERENCES

A. Fazeli and S. Houghten, ''Deep learning for the prediction of stock market trends,'' in Proc. IEEE Int. Conf. Big Data (Big Data), Dec. 2019, pp. 5513–5521.J.

B. Labiad, A. Berrado, and L. Benabbou, ''Short term prediction frame work for moroccan stock market using artificial neural networks,'' in Proc. 12th Int. Conf. Intell. Syst., Theories Appl. (SITA), 2018, pp. 1–6.

C.-Y. Lee and V.-W. Soo, ''Predict stock price with financial news based on recurrent convolutional neural networks,'' in Proc. Conf. Technol. Appl. Artif. Intell. (TAAI), Dec. 2017, pp. 160– 165.

K. A. Al-Thelaya, E.-S.-M. El-Alfy, and S. Mohammed, ''Forecasting of bahrain stock market with deep learning: Methodology and case study,'' in Proc. 8th Int. Conf. Modeling Simulation Appl. Optim. (ICMSAO), Apr. 2019, pp. 1–5

K. Khare, O. Darekar, P. Gupta, and V. Z. Attar, ''Short term stock price prediction using deep learning,'' in Proc. 2nd IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT), May 2017, pp. 482–486

M. Agrawal, A. U. Khan, and P. K. Shukla, ''Stock price prediction using technical indicators: A predictive model using optimal deep learning,'' Learning, vol. 6, no. 2, p. 7, 2019

P. Oncharoen and P. Vateekul, ''Deep learning for stock market prediction using event embedding and technical indicators,'' in Proc. 5th Int. Conf. Adv. Inform., Concept Theory Appl. (ICAICTA), Aug. 2018, pp. 19–24.

P. Oncharoen and P. Vateekul, ''Deep learning using risk reward functionfor stock market prediction,'' in Proc. 2nd Int. Conf. Comput. Sci. Artif.Intell. (CSAI), 2018, pp. 556–561.

P. Patil, C.-S.-M. Wu, K. Potika, and M. Orang, ''Stock market prediction using ensemble of graph theory, machine learning and deep learning mod els,'' in Proc. 3rd Int. Conf. Softw. Eng. Inf . Jan. 2020, pp. 85–92

pp. 51–61, Jul. 2018 pp. 74–78.

Q. Wang, W. Xu, and H. Zheng, ''Combining the wisdom of crowds and technical analysis for financial market prediction using deep random subspace ensembles,'' Neurocomputing, vol. 299,

S. Borovkova and I. Tsiamas, ''An ensemble of LSTM neural networks for high-frequency stock market classification,'' J. Forecasting, vol. 38, no. 6, pp. 600 619, 2019.

S. Demir, K. Mincev, K. Kok, and N. G. Paterakis, ''Introducing technical indicators to electricity price forecasting: A feature engineering study for linear, ensemble, and deep machine learning models,'' Appl. Sci., vol. 10, no. 1, p. 255, Dec. 2019

T. Matsubara, R. Akita, and K. Uehara, ''Stock price prediction by deep neural generative model of news articles,'' IEICE Trans. Inf. Syst., vol. 101, no. 4, pp. 901–908, 2018

W. Li and J. Liao, ''A comparative study on trend forecasting approach for stock price time series,'' in Proc. 11th IEEE Int. Conf. Anti-Counterfeiting, Secur., Identificat. (ASID), Oct. 2017,

X. Li, P. Wu, and W. Wang, ''Incorporating stock prices and news senti ments for stock market prediction: A case of Hong Kong,'' Inf. Process. Manage., vol. 57, no. 5, Sep. 2020, Art. no. 102212

Y. Qiu, H.-Y. Yang, S. Lu, and W. Chen, ''A novel hybrid model based on recurrent neural networks for stock market timing,'' Soft Comput., vol. 24, pp. 15273 15290, 2020

Y. Xu and V. Keselj, ''Stock prediction using deep learning and sentiment analysis,'' in Proc. IEEE Int. Conf. Big Data (Big Data), Dec. 2019, pp. 5573–5580.