Recent Methods in Stock Price Prediction: A Review

Yuxuan Li
College of Literature, Science and the Arts, University of Michigan, S State St, Ann Arbor, U.S.A.

Keywords: Stock Price Prediction, Neural Networks, Hybrid Models, ARIMA, BiCuDNNLSTM-1dCNN Model.

Abstract: This paper evaluates various methodologies for predicting stock prices, from traditional models such as the Autoregressive Integrated Moving Average (ARIMA) to more advanced neural networks. It addresses the inadequacies of the ARIMA model, particularly its limitations in predicting the future, which are commonplace in dynamic financial markets. Then, it introduces a hybrid model, the Bidirectional Cuda Deep Neural Network Long Short-Term Memory combined with a one-dimensional Convolutional Neural Network (BiCuDNNLSTM-1dCNN). This model excels at capturing both the long-term trends and short-term fluctuations essential for accurate financial forecasting. Through extensive preprocessing, the model ensures the highest quality of input data, leading to more reliable predictions. Comparative results demonstrate that the BiCuDNNLSTM-1dCNN model significantly surpasses both ARIMA and simpler neural networks in accuracy and reliability. The paper concludes with a call for continued advancement of hybrid modeling techniques to enhance the precision of forecasts and empower data-driven investment strategies in volatile markets.

1 INTRODUCTION

The stock market functions via a network of exchanges where shares of publicly traded corporations are transacted. Investors depend significantly on analytical insights to make prudent financial decisions (Billah, Sultana, Bhuiyan, & Kaosar 2024). Analyzing a company's performance through data is considered essential before making any investment. This emphasizes the significant role of data-driven strategies in helping investors evaluate a company's potential and make informed financial commitments. The Autoregressive Integrated Moving Average (ARIMA) model is a standard method for forecasting stock values. This approach integrates three components: autoregression (AR), which forecasts future prices using historical data; moving averages (MA), which smoothen fluctuations by factoring in past errors; and differencing (I), which stabilizes non-stationary data. While ARIMA provides accurate short-term predictions due to its simplicity, it struggles with non-linear patterns often present in financial markets. As described, the ARIMA model for stationary time series follows the structure, where future values are estimated using a combination of historical data points and previous forecasting errors (Hossain et al. 2018). Although effective for short-term predictions, ARIMA often fails to account for the complexities of fast-changing market trends. Although ARIMA offers reliable results for short-term forecasts, it often fails to capture the complexities of rapidly changing market trends. In recent years, neural networks have become prominent for their capacity to discern complex patterns and relationships in financial data. Researchers (Zhao, Hu, Liu, Lan, & Zhang, 2023) show that analyzing historical stock prices through recurrent neural networks (RNNs) or their variants improves forecasting accuracy. These models partition data into intervals, allowing them to uncover meaningful patterns and predict future trends. Given the growing reliance on neural networks for stock price forecasting, this paper provides a detailed review of various neural network models. It evaluates their performance in predicting stock prices and explores ways to enhance forecasting precision through hybrid architectures, such as the BiCuDNNLSTM-1dCNN model.

Li and Y.
Recent Methods in Stock Price Prediction: A Review.
DOI: 10.5220/0013528200004619
In *Proceedings of the 2nd International Conference on Data Analysis and Machine Learning (DAML 2024)*, pages 552-556
ISBN: 978-989-758-754-2

2 EVALUATING NEURAL NETWORK ARCHITECTURES FOR STOCK FORECASTING

2.1 CNN in Stock Price Forecasting

Convolutional Neural Networks (CNNs) are a category of deep learning models designed to examine structured inputs, such as photographs, by emulating the activities of the visual cortex in animals. These networks autonomously acquire and identify spatial patterns, varying from simple to intricate. A CNN has three fundamental types of layers: convolutional, pooling, and fully connected layers. The convolutional layers perform specialized linear operations to identify features while pooling layers help reduce dimensionality. Fully linked layers translate the gathered features into outputs, including class predictions (Yamashita, Nishio, Do, & Togashi, 2018).

When applied to stock price forecasting, CNNs have demonstrated effectiveness in extracting localized features from time-series data. In advanced models like BiCuDNNLSTM-1dCNN, CNNs play a vital role by identifying patterns within the data, and facilitating accurate predictions (Selvin et al., 2017). Nonetheless, despite their efficacy in capturing short-term dependencies, CNNs are limited in modeling long-term trends. The challenge has resulted in the development of hybrid approaches, such as the CNN-LSTM model, which integrates the feature extraction capabilities of CNNs with the temporal dependency modeling of Long Short-Term Memory (LSTM) networks to proficiently capture both short- and long-term dependencies (Hiransha et al., 2018).

2.2 RNN and Its Role in Sequential Data Analysis

RNNs are designed for the processing of sequential input through a loop structure that retains information from previous phases. These networks perform repetitive computations on each element within a sequence, where the outcome at each step depends on both the current input and the results from previous steps (Shiri, Perumal, Mustapha, & Mohamed, 2023). These abilities make RNNs useful in forecasting stock prices, as they can identify temporal patterns and correlations in financial data, helping the model analyze how historical prices influence future trends.

However, basic RNNs face challenges related to limited memory retention, which reduces their effectiveness when processing long sequences. This shortcoming, known as short-term memory, arises from disappearing or exploding gradient issues during training, especially with extensive datasets. As a result, while RNNs are well-suited for short-term predictions, they struggle with capturing complex dependencies over longer periods—an essential requirement for accurate stock forecasting. To alleviate these limitations, sophisticated models like LSTM networks and Gated Recurrent Units (GRUs) were created. These models incorporate memory mechanisms that allow the network to save and utilize essential information over prolonged sequences, hence enhancing their effectiveness in making financial predictions.

2.3 Advancements with LSTM Networks

LSTM networks are widely used in deep learning models, notably acknowledged for their capability to capture long-term dependencies in serial data, such as stock prices. LSTM models have been prevalent for financial prediction because of their ability to capture temporal correlations among data points (Chung & Shin, 2018). Their ability to remember pertinent information facilitates jobs that necessitate the examination of historical trends across time. The Grey Wolf Optimization-Elman Neural Network (GWO-ENN) model references LSTM as one of the benchmark models in its evaluation of various stock forecasting techniques. While LSTM models generally achieve competitive results based on metrics like mean square error (MSE), the GWO-ENN model, in this study, exhibited superior performance for one-day-ahead forecasts (Chung & Shin, 2018).

Compared to simpler models like the Elman Neural Network (ENN), which performs well for short-term memory retention (Zheng, 2015), LSTMs offer a more advanced approach by incorporating cell states and gating mechanisms. These features enable the model to make decisions on whether to keep or discard information, ensuring that relevant data is preserved for future use. Hybrid models, such as GA-LSTM— which combines genetic algorithms with LSTM networks—further improve forecasting accuracy, outperforming traditional approaches in stock market predictions (Chung & Shin, 2018).

LSTMs have proven particularly effective in predicting stock prices over extended periods, ranging from multiple days to even longer timeframes. Although ENN models are suitable for short-term tasks, such as forecasting the next day’s price, LSTM networks excel when it comes to capturing long-term trends, thanks to their sophisticated memory mechanisms

2.4 Incorporating GRU for Enhanced Efficiency

The Structure of the GRU Neural Network According to CEEMDAN-Wavelet is an advanced technique aimed at improving the precision of time series forecasting, (Qi, Ren, & Su 2023). Especially in the financial market, including stock index prediction. The model combines the CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) signal processing technique with the Wavelet Denoising method with the GRU neural network to make a financial forecast.

Initially, the original input noised data is deconstructed utilizing the CEEMDAN approach. CEEMDAN decomposes the signal into many frequency components known as Intrinsic Mode Functions (IMFs). The high-frequency intrinsic mode functions frequently encompass greater noise and demonstrate a diminished signal-to-noise ratio (SNR). Nevertheless, these high-frequency components may include significant short-term information that can affect the precision of forecasts. Wavelet denoising is utilized instead of outright discarding them. The wavelet thresholding method distinguishes valuable signal information from noise by examining the wavelet coefficients, which denote the intensity of various frequency components within the signal. By establishing a suitable threshold, the framework eliminates noise while preserving critical features, so ensuring that essential information is not compromised during the preprocessing phase.

After denoising the high-frequency components, the subsequent step is to amalgamate them with the intermediate and low-frequency IMFs, which generally exhibit reduced noise and encapsulate the long-term patterns of the data. This component combination process yields a reconstructed, denoised signal that embodies both the short-term and long-term attributes of the original data. Therefore, the Wavelet denoising method retains critical information across all frequencies, facilitating precise forecasting of the data.

During the second phase of the framework, the purified and denoised data is input into a GRU neural network for forecasting. GRUs are especially adept at processing time series data as they effectively learn long-term dependencies, managing sequential information across time while avoiding complications such as the vanishing gradient problem that can impact conventional RNNs. The GRU network manages the denoised data by regulating the retention of prior information via its update gate and the elimination of information through its reset gate. This enables the GRU to constantly equilibrate the retention of pertinent historical knowledge with the assimilation of fresh input to produce precise predictions regarding future patterns, such as stock prices.

The application of CEEMDAN and wavelet denoising guarantees that the GRU model is provided with superior input data. The framework enhances the forecasting accuracy of the GRU by efficiently distinguishing noise from the signal. This is especially beneficial in financial forecasting, where erratic data may result in imprecise projections. Furthermore, the GRU's more straightforward architecture relative to LSTMs enhances its computational efficiency, while yet preserving the capacity to catch both short-term variations and long-term trends in the data. The integration of sophisticated denoising methods and a robust predictive model renders the system exceptionally useful for time series forecasting, particularly in contexts such as stock market prediction, where data is frequently noisy yet holds significant information.

The CEEMDAN–wavelet model underwent thorough testing through ten trials on two prominent financial indices, the CSI300 and the S&P 500, to assess its predictive accuracy (Qi et al., 2023). The model's performance in these tests was evaluated using Mean Squared Error (MSE) and Mean Absolute Error (MAE), which are conventional metrics for measuring prediction accuracy. The model integrates CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) and wavelet denoising methods to preprocess the data, efficiently distinguishing valuable signal information from noise. The denoised data serves as input for a GRU neural network, which learns and forecasts future trends. Comparative evaluations were performed against other prominent models, including LSTM, GRU, CNN-BiLSTM, and ANN. Statistical research, including t-tests, demonstrated that the CEEMDAN–wavelet model consistently attained the lowest MSE and MAE values, exhibiting significant differences from the other models. The CEEMDAN–wavelet model successfully manages noisy data and captures intricate patterns, rendering it more accurate and dependable for time series forecasting compared to the classical and neural network models evaluated.

2.5 A Hybrid Model

The BiCuDNNLSTM-1dCNN model, presented by (Kanwal, Lau, Ng, Sim, & Chandrasekaran, 2022) commences with a crucial data preparation phase. Given the inconsistent nature of financial datasets, particular attention is given to handling missing values (NaNs). To maintain the integrity and continuity of the data stream, NaN values are replaced with the average of adjacent values (Kanwal et al., 2022). Furthermore, normalization is executed via the Min-Max scaler, which adjusts the feature values to a range of 0 to 1, thus enhancing the model's performance and computational efficiency. Once preprocessing is complete, the data is divided into 90% for model training and 10% for testing the predicted accuracy.

The configuration of hyperparameters is the initial phase in the complex training process of BiCuDNNLSTM-1dCNN. A Dropout layer is utilized to eliminate nodes at a rate of 20% during training to mitigate overfitting, alongside a bidirectional CuDNNLSTM layer that proficiently collects features from sequential input, followed by a one-dimensional CNN layer that identifies sudden market swings, crucial for capturing quick changes in the stock market. (Kanwal et al., 2022). Hyperparameter tuning is performed with a random search strategy, which navigates the parameter space more effectively than conventional grid search methods, resulting in decreased computational time and enhanced model performance (Bergstra & Bengio, 2012).

The practical execution of the model commences with the preparation of stock market data, encompassing open, maximum, minimum, and closing prices, along with trading volume, to ensure uniform inputs for training. The model uses a lookback window of 50-time steps to examine historical patterns, enabling the BiCuDNNLSTM component to leverage its bidirectional functionality effectively.

The BiCuDNNLSTM-1dCNN model has consistently demonstrated superior performance across various datasets, including individual stocks like Crude Oil and DAX ETF, as well as major indices such as GDAXI and HSI. It outperforms other models, including CNN, RNN, and standard LSTM, with lower RMSE and MAE values, indicating enhanced accuracy (Kanwal et al., 2022). This model functions as a dependable instrument for financial analysts and traders, demonstrating the efficacy of sophisticated machine-learning methodologies in intricate financial forecasting endeavors. By integrating LSTM and CNN architectures, the model effectively combines temporal sequence analysis with pattern recognition, establishing BiCuDNNLSTM-1dCNN as a robust solution for stock market prediction in volatile environments.

3 CONCLUSIONS

Forecasting stock prices has always presented a difficulty due to the inherent complexity and volatility of trade markets. Although conventional methods such as ARIMA are proficient for short-term forecasts, they frequently fail to account for the nonlinear dynamics inherent in stock data. Conversely, neural networks—such as CNNs, RNNs, LSTMs, and GRUs—have arisen as formidable instruments for stock price prediction, delivering superior performance through modeling intricate temporal dependencies and nonlinearities. The creation of hybrid architectures, shown by the BiLSTM-CNN model, enhances prediction accuracy and stability by integrating the advantages of various neural network models.

CNNs excel at recognizing limited, transient patterns in time series data, rendering them effective for forecasting rapid market fluctuations. Nonetheless, CNNs independently may encounter difficulties in capturing long-range relationships, hence constraining their utility in contexts necessitating extensive trend analysis. RNNs, especially LSTMs, and BiLSTMs, excel in handling sequential dependencies by retaining crucial information over extended periods, making them popular choices for financial forecasting. GRUs provide a more efficient solution by streamlining the architecture, diminishing computing complexity, and accelerating training without compromising accuracy, rendering them suitable for situations necessitating rapid, dependable predictions.

In conclusion, neural network models, particularly hybrid systems that amalgamate various designs, have continuously surpassed traditional forecasting methods. Their capacity to comprehend intricate patterns, retain vital information over time, and adeptly handle sequential dependencies enables them to deliver enhanced accuracy and reliability, which is crucial for investors making data-driven judgments in the dynamic and unpredictable realm of finance. As financial markets progress, the demand for advanced forecasting techniques will increase. Consequently, continuous research and development in machine learning, especially in the creation and enhancement of hybrid neural network models, will be crucial. These models can provide profound insights, adjust to fluctuating market conditions, and deliver more accurate predictions, allowing investors to make informed decisions that enhance their returns. Utilizing these improvements, the future of stock price forecasting will probably experience enhanced incorporation of artificial intelligence, promoting data-driven investment techniques that are resilient and adaptable to market complexity.

REFERENCES

Billah, M. M., Sultana, A., Bhuiyan, F., & Kaosar, M. G. 2024. Paper templates. *Stock price prediction: comparison of different moving average techniques using deep learning model*. NEURAL COMPUTING AND APPLICATIONS, 36(11), 5861-5871.

Kanwal, A., Lau, M. F., Ng, S. P., Sim, K. Y., & Chandrasekaran, S. (2022). *BiCuDNNLSTM-1dCNN—A hybrid deep learning-based predictive model for stock price prediction*. EXPERT SYSTEMS WITH APPLICATIONS, 202, 117123.

Kumar Chandar, S. (2021). *Grey Wolf optimization-Elman neural network model for stock price prediction*. SOFT COMPUTING, 25, 649-658.

Ma, Q. (2020). *Comparison of ARIMA, ANN and LSTM for stock price prediction*. IN E3S WEB OF CONFERENCES (VOL. 218, P. 01026). EDP SCIENCES.

Qi, C., Ren, J., & Su, J. (2023). *GRU neural network based on CEEMDAN–wavelet for stock price prediction*. APPLIED SCIENCES, 13(12), 7104.

Shiri, F. M., Perumal, T., Mustapha, N., & Mohamed, R. (2023). *A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU.* ARXIV PREPRINT ARXIV:2305.17473.

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). *Convolutional neural networks: an overview and application in radiology.* INSIGHTS INTO IMAGING, 9, 611-629.

Zhao, C., Hu, P., Liu, X., Lan, X., & Zhang, H. (2023). *Stock market analysis using time series relational models for stock price prediction.* MATHEMATICS, 11(5), 1130.