A Hybrid Stock Price Prediction Model Based on PRE and Deep Neural Network

Srivinay 1,*,†, B. C. Manujakshi 1,†, Mohan Govindsa Kabadi 2,† and Nagaraj Naik 3,*,†

1 Department of Computer Science and Engineering, Presidency University, Bangalore 560065, India; manujakshi@presidencyuniversity.in
2 Department of Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to Be University), Bangalore 561203, India; mkabadi@gitam.edu
3 Nitte Meenakshi Institute of Technology, Bangalore 560064, India
* Correspondence: srivinay@presidencyuniversity.in (S.); nagaraj.naik@nmit.ac.in or nagara21.naik@gmail.com (N.N.)
† These authors contributed equally to this work.

Abstract: Stock prices are volatile due to different factors that are involved in the stock market, such as geopolitical tension, company earnings, and commodity prices, affecting stock price. Sometimes stock prices react to domestic uncertainty such as reserve bank policy, government policy, inflation, and global market uncertainty. The volatility estimation of stock is one of the challenging tasks for traders. Accurate prediction of stock price helps investors to reduce the risk in portfolio or investment. Stock prices are nonlinear. To deal with nonlinearity in data, we propose a hybrid stock prediction model using the prediction rule ensembles (PRE) technique and deep neural network (DNN). First, stock technical indicators are considered to identify the uptrend in stock prices. We considered moving average technical indicators: moving average 20 days, moving average 50 days, and moving average 200 days. Second, using the PRE technique-computed different rules for stock prediction, we selected the rules with the lowest root mean square error (RMSE) score. Third, the three-layer DNN is considered for stock prediction. We have fine-tuned the hyperparameters of DNN, such as the number of layers, learning rate, neurons, and number of epochs in the model. Fourth, the average results of the PRE and DNN prediction model are combined. The hybrid stock prediction model results are computed using the mean absolute error (MAE) and RMSE metric. The performance of the hybrid stock prediction model is better than the single prediction model, namely DNN and ANN, with a 5% to 7% improvement in RMSE score. The Indian stock price data are considered for the work.

Keywords: prediction rule ensembles; deep neural network; moving average

1. Introduction

The financial stock market is dynamic [1,2]. Stock price trading is riskier due to high volatility [3]. The volatility estimation in the financial market is a challenging task. The financial market is volatile due to many internal and external factors [4]. Internal factors such as company earnings, political issues, and reserve bank policy affect stock prices. External factors, such as geopolitical tension, e.g., Russia and Ukraine, COVID-19, oil prices, and commodity market ups and downs, impact stock prices. Due to high volatility and nonlinearity in stock prices, an accurate stock predicting model is one of the hot research fields. An accurate prediction model helps investors to gain profits from the financial market.

In recent years, technical features [5] and fundamentals data [6] have been considered to identify the trend in stock prices. Here, trends can be classified as uptrend and downtrend in stock prices. Technical indicators are computed based on a mathematical formula [7].
These technical indicators are helpful to identify the trends in stock prices. Therefore, in this work, we considered technical indicators to identify the trends in stock prices.

Patel et al. [8] considered technical indicators to identify the trend in stock price. The study stated that if technical indicator relative strength index (RSI) value is greater than 70, the trend is down due to overbought stock prices. If technical indicator RSI value is less than 30, the trend is up due to oversold stock prices. In recent years, technical indicators were considered to identify the trend in stock prices. These technical indicators are trained in machine learning models to predict the future stock prices [9]. Technical indicators such as moving average, Bollinger bands, RSI, and MACD are commonly used.

An artificial-neural-network (ANN) [10]-based model was considered for stock price forecasting. ANN was inspired by a biological neural network, where each neuron accepts the input as stock price data and performs the task. The output of neurons is computed using nonlinear functions such as sigmoid and rectifier units. A support-vector-machine (SVM) [11]-based model was considered for stock price prediction. SVM is the most popular method for stock price classification. Stock prices are classified into two classes. One is up class, and the second is down class. In SVM, stock price data are constructed as a stock(s)-dimensional vector, and these dimensions are separated using a hyperplane. Since stock price data are nonlinear to classify classes, kernel tricks were used, such as sigmoid and Gaussian radial function. A random-forest (RF) [12]-based model was considered for stock prediction and classification. Stock price data are considered to construct the multiple decision trees; the decision tree aims to reduce variance in stock data. The average prediction of each decision tree is computed and selects the decision tree which has the lowest RMSE score.

A hybrid neural network VMD-LSTM [13] was considered to predict stock price indices. The VMD algorithm divides the original time series into subseries, and LSTM builds a training and prediction model for each subseries. The sum of all estimated subseries results was combined and considered as final prediction results. Jing et al. [14] proposed a hybrid model that combines deep learning and sentiment analysis to forecast the Chinese stock market. The hybrid model integrates LSTM neural networks for stock prediction with convolutional neural networks for sentiment analysis. Senapati et al. [15] proposed a hybrid model for stock price prediction using ADALINE neural networks and modified particle swarm optimization (PSO). It uses PSO to optimize and update weights of ADALINE neural networks. Kim et al. [16] proposed a hybrid long short-term memory (LSTM) model that combines LSTM and GARCH-type models to forecast stock price volatility. The LSTM model is an RNN variation; it creates a self-loop to generate a gradient path that can flow continuously for a long time. Each iteration changes the self-weight loops, preventing gradient vanishing when the RNN model updates the weights. The hybrid LSTM model performance was better than GARCH models. DNN [17,18] can solve nonlinear problems better than conventional machine learning techniques. A DNN-based prediction model is developed using the PSR approach and long- and short-term memory networks (LSTMs) for DL, and it is used to predict stock price movements. Sedighi et al. [19] proposed a hybrid stock prediction model by combining ABC-ANFIS-SVM. To optimize the technical indicators, they considered the ABC-ANFIS combination in the SVM forecasting model.

Hu et al. [20] proposed combined deep learning and the GARCH model to predict copper price volatility. One ANN is connected with two RNNs (LSTM and BLSTM) to create hybrid neural networks. Zhong et al. [21] proposed a DNN-based method for stock price prediction. The performance of DNN depends on the data representation format and the network structure, activation function, and model parameters. A 16 support vector regression model is considered for stock forecasting. The grid search method was used to find the optimal kernel function and fine-tune the SVM parameters to improve the model prediction accuracy.

Most of the work considered the hybrid models for stock price prediction to deal with the nonlinear data and improve the prediction accuracy in the model. The combination
of PRE and DNN methods has not been studied in the literature. The reason is that both models are very robust for nonlinear data.

The contribution of this work is as follows:

(1) First, stock technical indicators are considered to identify the uptrend in stock prices. We consider moving average technical indicators: moving average 20 days, moving average 50 days, and moving average 200 days.

(2) We propose a hybrid stock prediction model using the PRE and DNN.

The paper’s organization is as follows: Section 2 describes the literature review, and Section 3 data specification. Section 4 presents the moving average to identify an uptrend in stock price. Section 5 presents a hybrid stock price prediction model. Section 6 describes the results, and Section 7 concludes the work.

2. Literature Reviews

A CNN-BiLSTM-AM [22] approach was considered to forecast the stock price for the next day based on historical stock price data. Opening prices, highest and lowest prices, closing prices, volume, and turnover are all inputs that are used in this method. Ups and downs are also taken into account. The input data is processed by a convolutional neural network (CNN). The extracted feature data is learned and predicted using BiLSTM. Using AM, one could see how various periods in a time series dataset impact prediction results. Experiment results using CNN-BiLSTM-AM were the most accurate and had the best prediction accuracy compared to the other models.

ANN [23] was used to forecast the stock’s closing price for the next day, and RF was also used for comparison analysis. Based on RMSE, MAPE, and MBE values, it is also evident that ANN outperforms RF when it comes to stock price forecasting. Experimental results demonstrate that ANN performed better than other methods with RMSE (0.42), MAPE (0.77), and MBE (0.013).

GWO-optimized ENN [24] was considered to improve stock forecasting accuracy. The proposed model used the daily stock prices of eight businesses from the NASDAQ and the New York Stock Exchange to calculate daily returns. The ENN model was optimized by using the GWO algorithm. Results show that GWO-ENN outperforms the optimized models such as PSO-ANN and FPA-ANN.

The combination model ARI-MA-LS-SVM [25] was used in this work to forecast stock prices for the domestic stock market. The input variables are reduced in dimension with the principal component analysis method. Due to this, the model’s training time was lowered.

K-NN regression [26] was used to forecast the stock prices. The stock prices of various companies were studied and various technical indicators considered to predict the future movements of stock prices. The experimental results performed better compared to other machine learning approaches.

Zhang et al. proposed [27] an ensemble machine learning model called SVR-ENANFIS for stock forecasting by merging support vector regression with ensemble adaptive neuro-fuzzy inference. ANFIS is the first two-stage model to deal with deviations between technical indicator forecast results on the (t + n)th day and the actual value on that day, which is included. Furthermore, GA estimates the optimal input length and the optimal SVR hyperparameters to improve technical indicator prediction accuracy.

Xu et al. [28] considered the clustering method in the two-stage prediction model. It consists of technical indicators and stock price data. A two-stage prediction model employed an ensemble learning algorithm with clustering techniques. The results of the experiments reveal that for financial stock forecasting, the hybrid prediction model has the best overall prediction accuracy.

A deep reinforcement learning strategy for stock transactions was implemented [29], which demonstrated the usefulness of DRL, and it helps to identify the financial strategy for stock trading. It also compared three different traditional DRL models. The DQN model had the best results when dealing with stock market strategy decision-making problems.
A DNN-based LSTM network model [17] and the PSR time series analysis approach were combined. Preprocessing involves denoising and normalizing the data. Partitioning with a temporal window was used for data structures. The different activation function was considered in the DNN model. The results are compared with the existing conventional SVR machine learning method, a deep MLP model, a deep LST model without the PSR process, and a deep LST model paired with PSR. It was found that the DNN-based LSTM model was more accurate than the others.

Prediction’s study [30] shows that the proposed feature selection method comprises feature set creation and identification of a relevant feature set, which is responsible for effective performance in stock price predictions. The initial model considered 68 technical indicators. A total of 94% accurate predictions were made using the KOSPI dataset’s expanded feature set 2, which increased accuracy by 33% from the original feature set’s accuracy of 61% and by 18% from the expanded feature set’s accuracy of 75%, which was used to replicate a benchmark study using commonly used technical indicators.

Jing et al. [14] considered a convolutional neural network model to classify the investor’s hidden attitudes from a large stock forum. The hybrid model was proposed using the LSTM neural network approach for assessing stock market technical indicators and the sentiment analysis results performance. In terms of F-measures, the sentiment analysis model was outperformed. The Shanghai Stock Exchange’s thirty equities were considered to build the prediction model (SSE). The hybrid model has a reduced MAPE value of 0.0449, outperforming the single model. Using LSTM neural networks, it was found that combining investor sentiment with technical indicators improves the accuracy of future stock price predictions. The FinALBERT [31] model has been considered for stock price classification.

Zhang et al. [32] considered technical indicators to extract the trading rules for stock price prediction. HAR-PSO-ESN-based reservoir computing was applied [3] in combination with the established HAR recurrent neural network to create a new HAR-PSO-ESN model. The ESN architecture’s hybrid model was used as the PSO metaheuristic to adapt its hyperparameters based on the HAR’s handcrafted features. In most cases, the model anticipates more accurately than other models. The related work is depicted in Table 1.

**Table 1. Related works.**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Technique</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niu et al. [13]</td>
<td>Hybrid neural network VMD-LSTM</td>
<td>Stock price prediction</td>
</tr>
<tr>
<td>Jing et al. [14]</td>
<td>Hybrid model that combines deep learning and sentiment analysis</td>
<td>Stock price prediction</td>
</tr>
<tr>
<td>Senapati et al. [15]</td>
<td>Hybrid model for stock price prediction using ANN and PSO</td>
<td>Stock price prediction</td>
</tr>
<tr>
<td>Kim et al. [16]</td>
<td>Hybrid long short-term memory (LSTM)</td>
<td>Stock price forecasting</td>
</tr>
<tr>
<td>Sedighi et al. [19]</td>
<td>Hybrid stock prediction model by combining ABC, ANFIS, and SVM</td>
<td>Stock price forecasting</td>
</tr>
</tbody>
</table>

### 3. Data Specification

The Indian stock market [33] data are used in the experiment. The stock data contain daily stock price information such as open, high, low, close, and volume. In total, 4500 trading days are considered from 1 January 2007 to 10 October 2021. For the experiment, we considered the top nifty bank stock. The stock is Kotak, ICICI, Axis, and Yes Bank. The reason for selecting banking stock is due to high market capital. Hence, it is more volatile. The overall work is described in Figure 1. First, stock technical indicators are considered to identify the uptrend in stock prices. We considered moving average technical indicators. Second, using the PRE technique, we computed different rules for stock prediction. Third,
the three-layer DNN is considered for stock prediction. Fourth, the average results of the PRE and DNN prediction model are combined. The code is released in the supplementary.

**Figure 1. Overall work.**

4. **Stock Uptrend Identification Using Moving Average Technical Indicators**

Technical indicator [34] is a popular statistical technique to identify the trend in the stock. Here, the trend can be an uptrend or downtrend. In this work, we considered moving average technical indicators to identify the trend in the stock. The advantage of moving average technical indicators is that it reduces the random behavior of the stock price. However, most of the work considered a single moving average technical indicator to identify an uptrend in the stock. Identifying the accurate uptrend in stock required the combination of moving average technical indicators for uptrend confirmation. Therefore, we considered three moving average technical indicators to identify the uptrend in the stock price. Technical indicators are computed for 20 days, 50 days, and 200 days, based on the formula of [35]. Stock trading above the 20 days, 50 days, and 200 days moving average is considered an uptrend in stock; otherwise, it is a downtrend in stock price. The uptrends of stock prices are given as input to the hybrid stock prediction model for future stock prediction.

5. **Hybrid Stock Prediction Model Using Prediction Rule Ensembles and Deep Neural Network**

The hybrid prediction model is the most popular method for stock price prediction, which produces superior results [36,37]. The hybrid-based prediction learning model has emerged as one of the most effective learning approaches for stock price prediction. Due to their bias in variable selection and instability [38,39] they cannot be used effectively. This means that small changes in the training data can significantly impact the final results. Therefore, to overcome instability in the prediction model, we proposed hybrid stock prediction based on PRE and DNN. A hybrid stock prediction model is depicted in Figure 2. Later, the prediction rule ensembles method and the DNN method are combined to enhance the prediction results.
5.1. Prediction Rule Ensembles

PRE is a sparse collection of rules that generate different decision trees, which is helpful for classification and regression-based problems [40]. The stock uptrend data are given as input to PRE, creating the different combinations of trees. Each prediction rule is evaluated based on logical statements, i.e., if and else conditions. Each node in a tree can be represented as 0 and 1 rule, and it is defined in Equation (1), where K represents the total number of trees, \( F(p) \) represents the prediction function, and \( x_0 \) represents the weight of the tree. We selected the decision tree with the lowest RMSE score. The different prediction rules generated by PRE methods are depicted in Figures 3–5.

\[
F(p) = x_0 + \sum_{k=1}^{K} x_k F_k(p)
\]  

(1)
5.2. DNN for Stock Price Prediction

DNN is most commonly used to deal with nonlinear data [17]. In the proposed work, we considered the three-layer DNN model. First is the input layer, which is composed of a set of neurons. Each neuron in the input layer represents the stock price data. Here, open price, low price, high price, and volume price, technical indicator features, are given as input to the first layer of the DNN model. The weight is adjusted to input data for learning purposes. In Equation (2), $W$ denotes weight, $B$ denotes bias, $H$ is the hidden layer, and $\delta$ denotes the activation function. The outcome of each neuron in the first layer is computed using the activation function; this acts as a nonlinear function in the model. The second layer is the hidden layer, and the third layer is the output layer. The activation function is used to activate the neurons in the DNN. The rectifier linear unit and the hyperbolic tangent (tanh) are two activation functions that are considered in DNN architecture. Backpropagation computes the gradient of the loss function for the DNN model and adjusts weight accordingly. We have fine-tuned the hyperparameters manually for better accuracy, such as the number of layers, learning rate, neurons, and number of epochs in the model.

The hybrid stock price prediction algorithm steps are described in Algorithm 1.

**Algorithm 1** Hybrid stock prediction model using PRE and DNN.

1: Input stock price data from NSE portal.
2: Compute the Moving Average for 20, 50 and 200 days of stock price.
3: \textbf{Moving Average} = Stock Price1 + Stock Price2 + … Stock PriceN / Total Number of Days
4: Combine the moving average to identify the uptrend in stock price.
5: if (Stock Price > 20 Days & Stock Price > 50 Days & Stock Price > 200 Days)
6: then Uptrend in Stock Price
7: Else
8: Down Trend in Stock Price
9: Uptrend Stock Price are given input to PRE method.
10: Rule are generated using prediction rule ensembles method.
11: $F(p) = x_0 + \sum_{k=1}^{K} x_k F_k(p)$
12: Selected the decision tree with the lowest RMSE score.
13: Later, Uptrend stock data given input to the DNN.
14: Constructed Three Layer DNN model.
15: Compute $H = \delta(WH_1 + B)$
16: Fine-tune the hyperparameters of the DNN method, such as the number of layers, learning rate, neurons, and number epoch in the model.
17: Average results of the PRE and DNN prediction model are combined.
18: Validate the results using 10 cross fold validation.
19: Evaluate the performance of model using RMSE and MAE metric.
\[ H = \delta(WH_1 + B) \]  

(2)

6. Results

For the experimental work, we considered Indian stock exchange data which is publicly available at https://www.nseindia.com/ (accessed on 16 April 2022). The datasets comprise daily stock price data for 4500 trading days covering from 1 January 2007 through 10 October 2021. Stock technical indicators are considered to identify the uptrend in stock prices using moving average technical indicators. The proposed hybrid stock prediction model is evaluated using MAE and RMSE, an equations are defined below (3) and (4). The observed value is represented by the variable \( y_i \). The predicted value is defined by variable \( x_i \), while the number of elements in the dataset is represented by the variable \( n \). To obtain the best results, we fine-tuned the hyperparameters of DNN, and they are depicted in Table 2. The training loss and testing loss are computed, and they are depicted in Figures 6–8. The objective of the loss function is to reduce the error rate in the model; for that, we increased the number of epochs, and we found that after 450 epochs, the model starts to converge. The proposed hybrid stock price prediction model performed better than DNN and ANN models, and it is depicted in Table 3.

\[
mae = \left( \frac{1}{n} \right) \sum_{i=1}^{n} |y_i - x_i|
\]  

(3)

\[
rmse = \sqrt{\left( \frac{1}{n} \right) \sum_{i=1}^{n} (y_i - x_i)^2}
\]  

(4)

Table 2. Hyperparameters.

<table>
<thead>
<tr>
<th>HyperParameters</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of layers</td>
<td>2 to 3</td>
</tr>
<tr>
<td>Number of neurons</td>
<td>5 to 20</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.001 to 0.004</td>
</tr>
<tr>
<td>Epochs</td>
<td>300 to 600</td>
</tr>
</tbody>
</table>

Table 3. Results.

<table>
<thead>
<tr>
<th>Stock</th>
<th>Prediction Technique</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kotak Bank</td>
<td>Proposed Hybrid Model</td>
<td>9.50</td>
<td>6.50</td>
</tr>
<tr>
<td>ICICI Bank</td>
<td>Proposed Hybrid Model</td>
<td>10.30</td>
<td>5.60</td>
</tr>
<tr>
<td>Axis Bank</td>
<td>Proposed Hybrid Model</td>
<td>9.90</td>
<td>7.10</td>
</tr>
<tr>
<td>SBI Bank</td>
<td>Proposed Hybrid Model</td>
<td>8.53</td>
<td>6.30</td>
</tr>
<tr>
<td>Infosys</td>
<td>Proposed Hybrid Model</td>
<td>9.30</td>
<td>7.20</td>
</tr>
<tr>
<td>TCS</td>
<td>Proposed Hybrid Model</td>
<td>8.55</td>
<td>6.75</td>
</tr>
<tr>
<td>Kotak Bank</td>
<td>DNN Model</td>
<td>13.50</td>
<td>11.50</td>
</tr>
<tr>
<td>ICICI Bank</td>
<td>DNN Model</td>
<td>14.50</td>
<td>13.60</td>
</tr>
<tr>
<td>Axis Bank</td>
<td>DNN Model</td>
<td>13.30</td>
<td>12.60</td>
</tr>
<tr>
<td>SBI Bank</td>
<td>DNN Model</td>
<td>13.50</td>
<td>11.20</td>
</tr>
<tr>
<td>Infosys</td>
<td>DNN Model</td>
<td>15.13</td>
<td>13.11</td>
</tr>
<tr>
<td>TCS</td>
<td>DNN Model</td>
<td>14.45</td>
<td>12.40</td>
</tr>
<tr>
<td>ICICI Bank</td>
<td>ANN Model [10]</td>
<td>15.1221</td>
<td>19.9444</td>
</tr>
<tr>
<td>SBI BANK</td>
<td>ANN Model [10]</td>
<td>17.4341</td>
<td>23.1585</td>
</tr>
</tbody>
</table>
7. Conclusions

Stock price forecasting is challenging due to noisy, dynamic, and nonlinear data in the stock market. The accurate prediction of stock prices helps investors increase profits in the financial market. Identifying the trend in the market is a challenging task. We proposed a hybrid stock prediction model by combining the PRE and DNN models. The proposed model overcomes the instability in the prediction model. We have fine-tuned the DNN hyperparameters manually for better accuracy, such as the model’s number of layers, learning rate, neurons, and number of epochs. The hybrid stock price model is evaluated using the metrics MAE and RMSE. The proposed hybrid stock prediction model results are compared with the DNN and ANN. The performance of the proposed model’s RMSE score is 5% to 7% lower than the existing DNN and ANN models. However, we considered limited technical indicators in the hybrid model. Exploring the different technical indicators alongside candlestick pattern identification can be future work.


Author Contributions: Methodology, S.; Validation, S.; Conceptualization, B.C.M., Supervision, B.C.M.; Conceptualization, M.G.K., Supervision, M.G.K.; Conceptualization, N.N.; Validation, N.N. All authors have read and agreed to the published version of the manuscript.

Funding: We would like to thank Presidency University Bangalore, GITAM University Bangalore Campus, and Nitte Meenakshi Institute of Technology, Bangalore, for their support for to carry the Research Work.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.
References


3. Ribeiro, G.T.; Santos, A.A.P.; Mariani, V.C.; dos Santos Coelho, L. Novel hybrid model based on echo state neural network applied to the prediction of stock price return volatility. *Expert Syst. Appl.* 2021, 184, 115490. [CrossRef]


