Intraday Stock Prediction Based on Deep Neural Network

Nagaraj Naik[1] · Biju R. Mohan[1]

Abstract Predicting stock price movements is difficult due to the speculative nature of the stock market. Accurate predictions of stock prices allow traders to increase their profits. Stock prices react when receiving new information. During the trading day, it is difficult to understand the up and down movements signaled by stock prices. This paper addresses the problem of fluctuations in stock prices. We proposed the method to identify stock movement trend in data, and this method considered the combination of candlestick data and technical indicator values. The outcome of this method is given as inputs to a deep neural network (DNN) to classify a stock price's up and down movements. National Stock Exchange, India, datasets are considered for an experiment from the years 2008 to 2018. The work is carried out using H2O deep learning on an RStudio platform. Experimental results are compared with a three-layer artificial neural network (ANN) model. The proposed five-layer DNN model outperforms state-of-the-art methods by 8–11% in predicting up and down movements of a given stock.

Keywords Technical indicators · DNN · Classification · Candlestick pattern

✉ Nagaraj Naik
  nagaraj21.naik@gmail.com

  Biju R. Mohan
  biju@nitk.edu.in

[1] Department of Information Technology, National Institute of Technology, Surathkal, Karnataka, India

Introduction

Due to the volatile nature of the stock market, predicting stock price movements for intraday trade is a challenging task. Interpretation of stock-related critical information makes trading more profitable [1]. The efficient market hypothesis (EMH) states that the prices of trade assets such as stocks already reflect all publicly available information, and thus, if your investments are based on publicly available information, you will not systematically outperform the market over time. Investors are making money by either taking a chance or making riskier investments [13]. However, most of the stock-related studies [4, 5] stated that the stock market can be predicted. Stock analysis can be carried out in two ways. The first is fundamental analysis, and the second is technical analysis. The fundamental analysis involves an examination of financial data, management, business track record, competition, earnings, macroeconomic factors, and overall economic conditions. The technical analysis is a method of making stock measurements based on historical price. This method involves various statistical analyses of stock data based on stock prices.

The extant literature suggests that the technical indicator is an essential parameter for evaluating stock prices [4, 7]. The first contribution of this paper is the identification of a trend in data by using a combination of candlestick data and technical indicators. The existing research has not considered these combinations. The second contribution of this paper is using DNN to classify and accurately predict a stock price's up and down movements. The existing research is based on a three-layer artificial neural network (ANN), which is unable to classify the up and down movements of stock prices.

Ranjeeta et al. [4] studied stock movement prediction on global indices, namely the Hang Seng index, Financial Times Stock Exchange, and Bombay Stock Exchange. The study considered the variational mode decomposition method for stock data, and machine learning methods were used to identify a nonlinear pattern in data. Paraskevi et al. [14] studied stock mid-price movement prediction using a limit order of stock price data. The work considered a set of features which include the prices, volumes, and ask and bid price of the stock. The performance of extracting features was evaluated based on a machine learning algorithm. Weng et al. [11] proposed multi-filter neural network method to extract stock price feature. The task considered stock-related features, namely open, close, high, low, and volume. Jing et al. [18] proposed sliding window method to cut raw the historical stock price data into the predefined length and proposed an unsupervised heuristic method for stock price prediction. Suryoday et al. [3] considered six technical indicators to predict the stock price using random forests and gradient boosted decision tree methods. Arash et al. [8] proposed a graph-based semi-supervised approach to predict up and down movements of stock prices. However, most of the existing work considered technical indicators to evaluate the stock price movements. Hsun [12] proposed candlestick-based approach for profitable trading in the stock market. However, most of the work considered technical indicators to evaluate the stock prices. Therefore, we have combined technical analysis and candlestick approach in this study.

A comparison between ANN and support vector machine (SVM) classifiers found that ANN performs better than SVM in forecasting the Istanbul Stock Exchange [7]. In another study, ten technical indicators were used to classify the up and down movements of stocks using ANN, SVM, random forest, and Naive Bayes, and it found that the random forest model outperformed the others [15]. The future performance of stock indices was studied using fuzzy time series modeling [16]. Zhang et al. [19] studied an optimization method with a backpropagation ANN to predict stock movements in prices. However, existing trading rules did not predict future periods when market conditions change dynamically. The augmented Dickey–Fuller test has been used to verify and assess the performance of trading rules [6]. Fei Wu et al. [17] discussed face attributes and considered different subspaces for learning the pattern. PCA-based feature extraction technique [9, 10] has been studied to recognize handwritten character, and SVM with different kernel parameters was considered for classification. Mallika et al. [2] proposed PCA method for facial emotion recognition and considered a random forest method for classification. Most of the work considered SVM and ANN classifier to predict the stock price up and down movements.

An overall framework for a stock prediction model described in Fig. 1. The following steps have been considered for extracting the technical indicators.

Step 1 Input National Stock Exchange (NSE) data
Step 2 Apply ten technical indicators [7]
Step 3 Calculate the technical indicators based on formulas [7]
Step 4 Output continuous technical indicator values

The following steps have been considered for extracting the candlestick pattern.

Step 1 Input NSE data
Step 2 Extract day low, high, open, and close price
Step 3 Apply Japanese's candlestick pattern on daily basis
Step 4 Output candlestick pattern

We have proposed a method to identify a trend in data using a combination of candlestick data and technical indicators, and it is described in Algorithm 1. In Algorithm 1, step 4, if SMA technical values are trading above the past 10-day average value of the stock price and candlestick pattern data are greater than the mean value of stock price, then it is considered as up movement, otherwise down movement (https://in.tradingview.com/ideas/candlestick/). In step 5, if EMA technical values are trading above the past 10-day average value of the stock price and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement. In step 6, if MOM technical values are trading above the past 10-day average value of the stock price and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement. In step 7, if
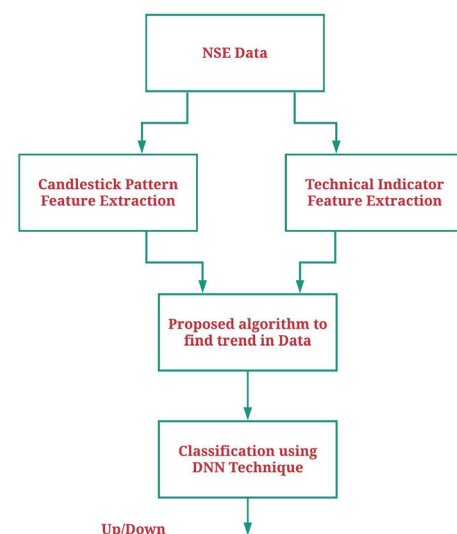

Fig. 1 Overall framework of proposed model

STCK technical values are trading above STCD 14-day average value of the stock price and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement.

In step 8, if RSI technical values are trading above 30 and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement. In step 9, if MACD technical values are trading above MACD 26-day average value of the stock price and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement. In step 10, if R technical values are trading below -80 and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement. In step 11 if A/D technical values are trading below -100 and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement. In step 12 if CCI technical values are trading above 100 and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement. In step 13 if EMA technical values are trading above past 20-day average value of the stock price and candlestick pattern data are greater than the mean value of the stock price, then it is considered as up movement, otherwise down movement.

The output of this algorithm is the modified technical indicators, which are inputs to the DNN prediction model.

DNN is one of the popular methods for learning lots of different types of pattern. The proposed model composed of more than one neural network layer. The neural network is designed based on human brains. Human brains consist of billions of small cells known as neurons. Kara et al. [7] considered a three-layer artificial neural network model for stock movements' classification, and it is described in Fig. 2. The proposed work, DNN, is considered to classify stock price's up and down movements, as shown in Fig. 3. It is five layers of architecture with three hidden layers, and each layer transfers the data between the neurons. It takes the input of ten technical indicator values, and it is denoted by x_i. Each of these technical indicators' values is multiplied by weight W_i, and it denoted as $W_i x_i$. Initially, some random values are assigned. Each of these weights is assigned to input value and weighted sum performed, and it is defined in Eq. (1). The bias is typical for each neuron. There is one bias per neuron. An activation function takes the weighted sum of input and bias. Kara et al. [7] study considered the sigmoid activation function. The sigmoid function is used for models where we have to predict the probability as an output. It exists between 0 and 1 and is represented by $f(x) = \frac{e^x}{e^x + 1}$. The ReLU activation function is considered in the experiment. ReLU is the most popular function used to deal with nonlinear data. It gives an output of x if x is positive and 0 otherwise, and it is represented as $f(x) = \max(0, x)$. There is no upper limit in ReLU. After training the DNN, it uses backpropagation method to improve the performance of the network. Cost function

---

Algorithm 1 Proposed algorithm to find trend in data

1: **Input :** Candlestick pattern and continues technical indicator values.
2: **Output :** Modified technical indicator values.
3: **for** each technical indicator i=1 to 10 **do**
4:     **if** (SMA >10 Days price AND Candlestick Pattern >Mean Value(Bullish Pattern)) **then** *T1=Up else T1=Down*
5:     **if** (EMA >10 Days price AND Candlestick Pattern >Mean Value) **then** *T2=Up else T2=Down*
6:     **if** (MOM >10 Days price AND Candlestick Pattern >Mean Value) **then** *T3=Up else T3=Down*
7:     **if** (STCK Fast Moving 14 DAYS >STCD Slow moving 14 days price AND Candlestick Pattern >Mean Value) **then** *T4=Up else T4=Down*
8:     **if** (RSI >30 price AND Candlestick Pattern >Mean Value) **then** *T5=Up else T5=Down*
9:     **if** (MACD Fast Moving 9 >MACD Slow Moving 26 days price AND Candlestick Pattern >Mean Value) **then** *T6=Up else T6=Down*
10:     **if** (R <-80 price AND Candlestick Pattern >Mean Value) **then** *T7=Up else T7=Down*
11:     **if** (A/D <-100 AND Candlestick Pattern >Mean Value) **then** *T8=Up else T8=Down*
12:     **if** (CCI >100 AND Candlestick Pattern >Mean Value) **then** *T9=Up else T9=Down*
13:     **if** (EMA >20 Days price AND Candlestick Pattern >Mean Value) **then** *T10=Up else T10=Down*
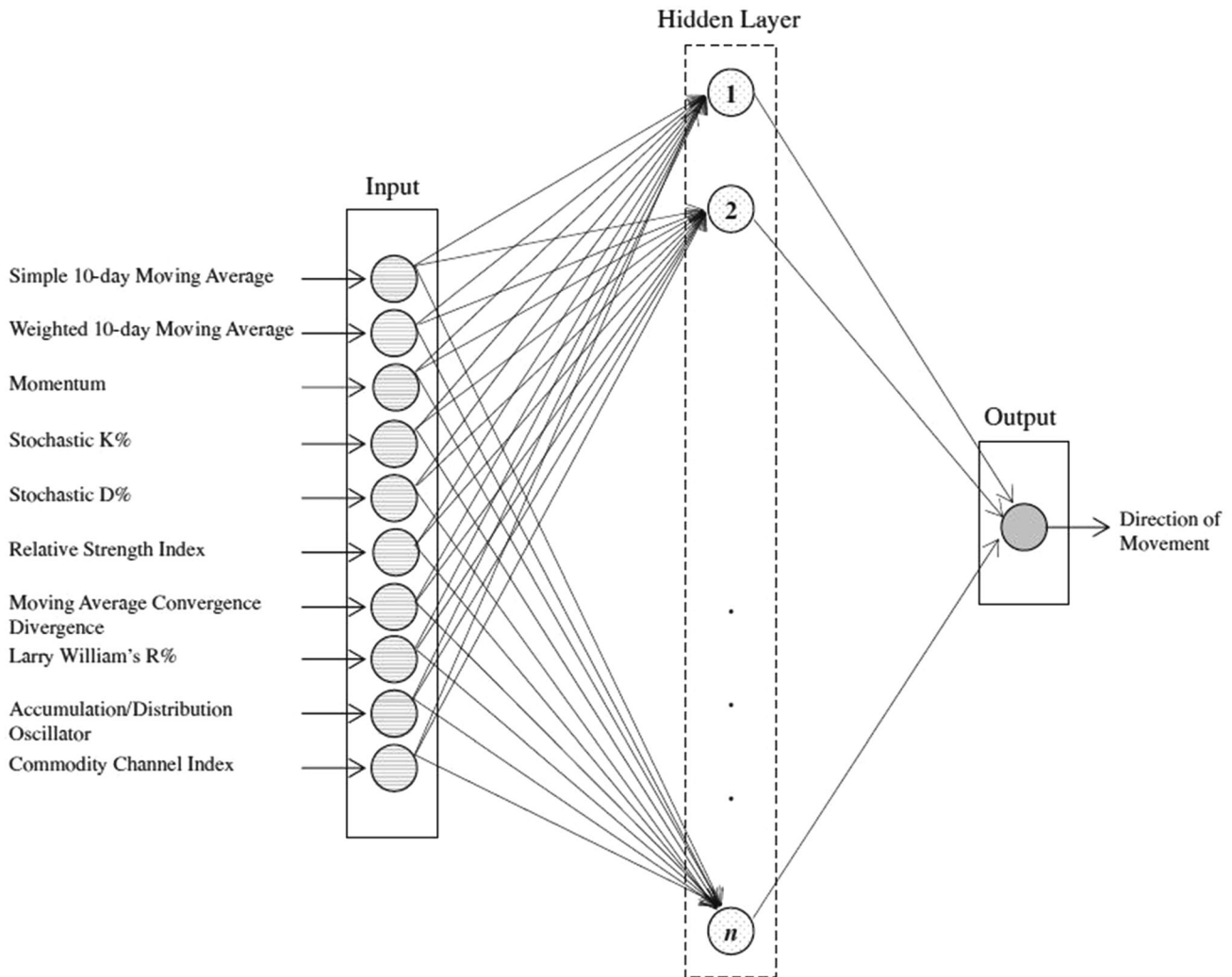14: **End for**.

Fig. 2 Three-layer ANN prediction model by Kara et al. [7]

helps to reduce the error rate. The cost value is the difference between DNN predicted output and the actual output. The least cost value is obtained by making adjustments to the weights and biases iteratively throughout the training process. Stochastic gradient descent (SGD) mechanism identifies how to change the weight and bias to minimize the cost functions. The DNN training process is nothing but finding the best value of weight and biases for each neuron in the network to maximize accuracy.

$$\alpha = \left(\sum_{i=1}^{n} W_i x_i + B \right) \tag{1}$$

In this work, stock price data are collected from NSE, India, and the website URL is (http://www.nseindia.com/stock). The stocks considered herein are Infosys, Reliance, HDFC, and HDFC Bank. The datasets range from the year 2008 to 2018. We have studied two commonly used metrics referenced in the literature, namely

accuracy and F measures, and they are described in Eqs. (2) and (3). These metrics are considered to evaluate the performance of DNN model. We use tenfold cross-validation in the experiment.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2}$$

$$F\text{Measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

Table 1 shows that the highest accuracies for Reliance stock, Infosys, HDFC Bank, and HDFC are 0.8361, 0.8178, 0.8211, and 0.8349, respectively. The performance of the proposed DNN model is better than the existing work and described in Table 1. The reason for the better performance is that the data preparation uses a combination of candlestick data and technical indicators. In addition, we have built a model on a five-layer deep neural network; the existing model was built on a three-layer artificial neural network.
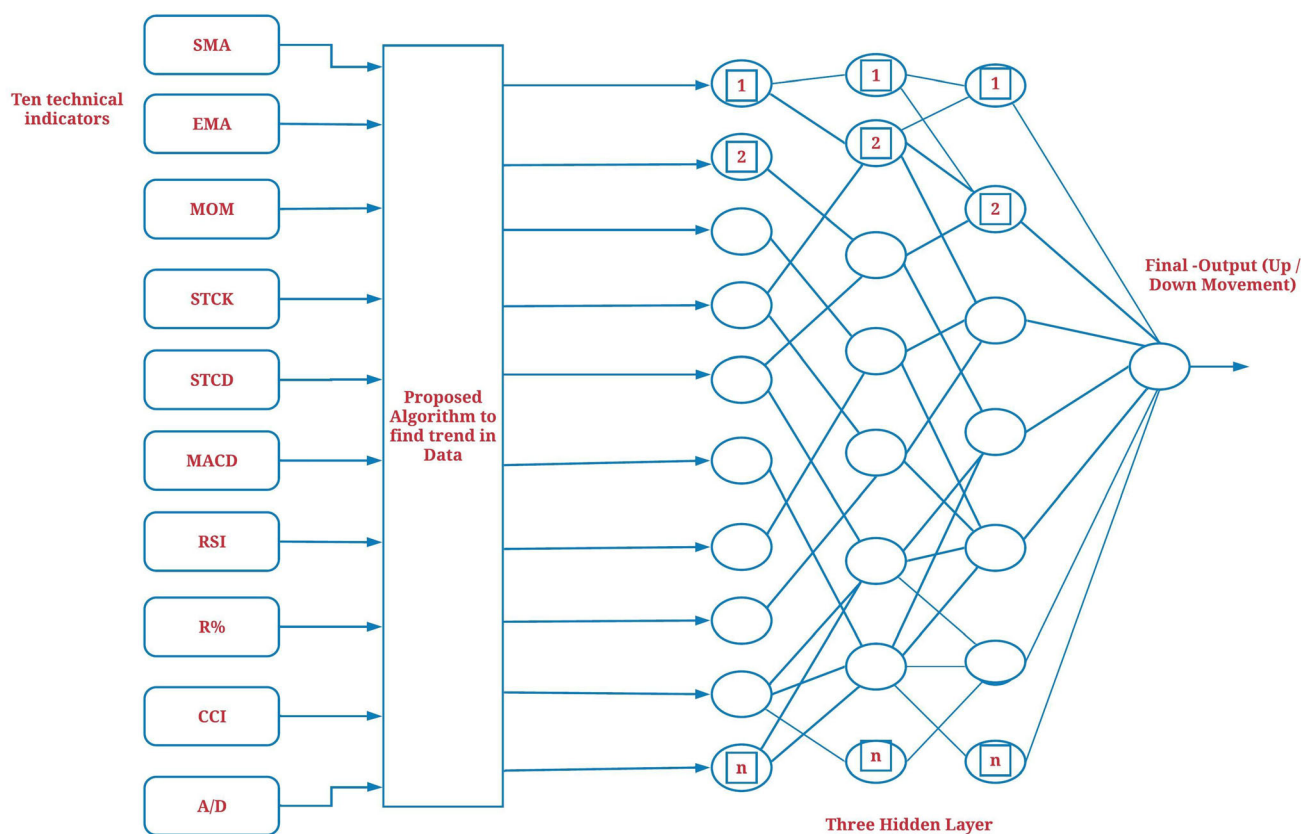
Fig. 3 Framework of proposed five-layer deep neural network

Table 1 Results' comparison with the existing work

Stock names	Prediction models	Accuracy	*F* measure
BSE	ANN [4]	0.74.94	0.7148
BSE	SVM [4]	0.76.79	0.7538
Reliance	ANN [7]	0.7275	0.7392
Reliance	Proposed model	0.8361	0.8646
Infosys	ANN [7]	0.7130	0.7364
Infosys	Proposed model	0.8178	0.8327
HDFC Bank	ANN[7]	0.7465	0.7315
HDFC Bank	Proposed model	0.8211	0.8410
HDFC	ANN[7]	0.7531	0.7432
HDFC	Proposed mode	0.8349	0.8558

In this study, we make predictions of the stock price's up and down movements in the NSE. The first contribution of this paper is the identification of a trend in data by using a combination of candlestick data and technical indicators. The existing research has not considered these combinations. The second contribution of this paper is using DNN to classify and accurately predict a stock price's up and down movements. The proposed deep neural network prediction model outperforms the ANN model. This paper focuses on intraday predictions. In the future work, we examine short-term and long-term prediction along with fundamental data analysis, such as a stock's price–earnings ratio, book value, and a company's quarterly profit.

References

1. Alexander SS (1961) Price movements in speculative markets: trends or random walks. Ind Manag Rev (pre-1986) 2(2):7
2. Arora M, Kumar M, Garg NK (2018) Facial emotion recognition system based on pca and gradient features. Natl Acad Sci Lett 41(6):365–368
3. Basak S, Kar S, Saha S, Khaidem L, Dey SR (2019) Predicting the direction of stock market prices using tree-based classifiers. N Am J Econ Finance 47:552–567
4. Bisoi R, Dash P, Parida A (2019) Hybrid variational mode decomposition and evolutionary robust kernel extreme learning machine for stock price and movement prediction on daily basis. Appl Soft Comput 74:652–678
5. Bollen J, Mao H, Zeng X (2011) Twitter mood predicts the stock market. J Comput Sci 2(1):1–8
6. Chen J-C, Zhou Y, Wang X (2018) Profitability of simple stationary technical trading rules with high-frequency data of Chinese index futures. Phys A Stat Mech Appl 492:1664–1678
7. Kara Y, Boyacioglu MA, Baykan ÖK (2011) Predicting direction of stock price index movement using artificial neural networks and support vector machines: the sample of the Istanbul stock exchange. Expert Syst Appl 38(5):5311–5319
8. Kia AN, Haratizadeh S, Shouraki SB (2018) A hybrid supervised semi-supervised graph-based model to predict one-day ahead movement of global stock markets and commodity prices. Expert Syst Appl 105:159–173
9. Kumar M, Jindal M, Sharma R (2014) A novel hierarchical technique for offline handwritten Gurmukhi character recognition. Natl Acad Sci Lett 37(6):567–572
10. Kumar M, Jindal M, Sharma R, Jindal SR (2018) Offline handwritten numeral recognition using combination of different feature extraction techniques. Natl Acad Sci Lett 41(1):29–33
11. Long W, Lu Z, Cui L (2019) Deep learning-based feature engineering for stock price movement prediction. Knowl Based Syst 164:163–173
12. Lu T-H (2018) Can profit targets improve the profitability of candlestick charting? Int J Revenue Manag 10(3–4):216–226
13. Malkiel BG, Fama EF (1970) Efficient capital markets: a review of theory and empirical work. J Finance 25(2):383–417
14. Nousi P, Tsantekidis A, Passalis N, Ntakaris A, Kanniainen J, Tefas A, Gabbouj M, Iosifidis A (2019) Machine learning for forecasting mid-price movements using limit order book data. IEEE Access 7:64722–64736
15. Patel J, Shah S, Thakkar P, Kotecha K (2015) Predicting stock market index using fusion of machine learning techniques. Expert Syst Appl 42(4):2162–2172
16. Rubio A, Bermúdez JD, Vercher E (2017) Improving stock index forecasts by using a new weighted fuzzy-trend time series method. Expert Syst Appl 76:12–20
17. Wu F, Jing X-Y (2017) Discriminant-reconstruction based multiple attribute classifiers for face recognition. Natl Acad Sci Lett 40(3):177–182
18. Zhang J, Cui S, Xu Y, Li Q, Li T (2018) A novel data-driven stock price trend prediction system. Expert Syst Appl 97:60–69
19. Zhang Y, Wu L (2009) Stock market prediction of s&p 500 via combination of improved bco approach and bp neural network. Expert Syst Appl 36(5):8849–8854