A Review: Machine Learning for Stock Market

Mashfiqur Rahman, Mohammad Nurul Huda

Department of Computer Science and Engineering, United International University Dhaka, Bangladesh

Corresponding author: Mashfiqur Rahman, Email: engg.mashfiqur@gmail.com

Financial markets are essential for a society's smooth operation. Because of its enormous financial impact, successfully anticipating stock market trends has long been a top priority for investors and scholars. Recently, a slew of studies have been undertaken for predicting stock market trends using deep learning algorithms, thanks to advances in deep learning algorithms, and their much-needed data and processing power. As a result, it's easy to get lost when looking for the correct references. We have summarized a few of the greatest papers in this area in this paper, which can serve as a guide for future researchers in the field.

Keywords: Stock Market, Deep Learning, Artificial Neural Network, Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), Convolutional Neural Network (CNN).

2021. In Raju Pal & Praveen K. Shukla (eds.), *SCRS Conference Proceedings on Intelligent Systems*, 305–312. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-93-91842-08-6-30

1 Introduction

Financial markets can be broadly referred as a marketplace where the trading of securities takes place. A stock represents the ownership of a fraction of a company [1]. The stock market is a marketplace or exchange where people may buy, sell, and issue shares of publicly traded corporations. [2]. It is also frequently referred as Capital Market, Equity Market, Stock Exchange etc. Stock trading can be dated back as far as the mid-1500s in Antwerp. However, the East India Company in London is credited for the modern stock trading [3]. Overtime, participants have tried different ways to predict the stock market trends. Ability of successfully predicting the trend would result in significant yield of profit. The prediction can be for trend prediction or the price prediction of the stocks. There are two major approaches when it comes to stock market predictions: 1) Fundamental Analysis, 2) Technical Analysis. Generally, fundamental analysis focus on a long-term of investing compared to the short-term approach preferred by technical analysis. However, it is understood that 90% of investors use technical analysis for stock market trend prediction.

Artificial intelligence (AI) is a branch of computer science which focus on making machines (computer programs) behave intelligently mimicking human intelligence. Machine Learning (ML) is a branch of AI. It works with computer algorithms that can make predictions or decisions by the use of data without being explicitly programmed [4]. In the last two decades, investors and researchers have looked into different ways to use machine learning for stock market prediction. Because the stock market's price is impacted by a variety of variables including national and international politics, natural catastrophes, religious events, human emotions, business policies, and so on, forecasting its price is extremely difficult owing to its nonlinear and complicated character. For these reasons, predicting stock market has gone through rigorous analysis by both academia and industry experts. Researchers have used time series data along with various heterogeneous market data and microeconomics variables. With increasing availability of the stock market data, computational capacity of the machines and advance in the field of machine learning, it is showing promising results for future success.

The rest of the paper is arranged as follows: Section 2 describes background materials for rest of the paper; Section 3 presents the reviewed paper with their methodology and results; Conclusion and future work is given in Section 4.

2 Background

2.1 Data

Machine learning methods learn from examples. These examples are known as data and the corresponding collection of data as dataset. The data can be mainly of two types:

1. Structured Data
2. Unstructured Data

These datasets are then divided into various subsets for training, validating and testing the deep learning models.

2.2 Artificial Neural Network-ANN

Artificial neural network, also referred as ANN or simply neural networks, is a computational algorithm which is inspired by the working principle of the brain. Neuron is the unit of an ANN. Based on the input to the network, different neurons either fires or keeps quiet. It is popular for its ability to adapt to changing inputs and find hidden patterns in the data. Figure 1 shows the architecture of a simple ANN.

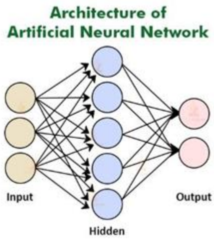

Fig. 1. Architecture of ANN [5]

The hidden units connect the input and output layers. The pattern of connection which comprises the number of layers between the input and output layers, number of nodes in each layer is known as the architecture of the neural network. There are two types of ANN architecture:

1. Single Layer Perceptron
2. Multi-Layer Perceptron

2.3 Multi-Layer Perceptron-MLP

Multi-Layer has at least three layers:

1. Input Layer,
2. Hidden Layer and
3. Output Layer.

Accept the input nodes, each node has a non-linear activation function. In an MLP, data travels from input to output layers in the forwards direction, similar to a feed forward network. In the MLP algorithm, the neurons are trained using the back propagation learning process. The main advantage of multi-layer perceptron is that it can distinguish data that is non-linear in nature. Error! Reference source not found and shows a simple MLP architecture with 2 hidden layers.

2.4 Recurrent Neural Network-RNN

Recurrent Neural Network, frequently known as RNN, which is a special type of ANN is generally used for sequential data modeling. They impact the current input and output by using the output from the previous phase. RNN's output is influenced by the sequence's earlier elements [6]. The internal architecture of a RNN is shown in Figure 2.

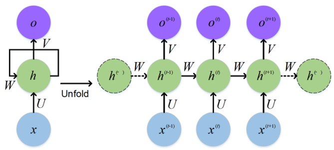

Fig. 2. Architecture of RNN [7]

2.5 Long Short-Term Memory-LSTM

LSTM was first introduced by Hochreiter et al. [8]. It is a variation of basic RNN algorithm. LSTMs overcome the vanishing and exploding gradients problem presented in RNN. Figure 3 shows basic architecture of LSTM.

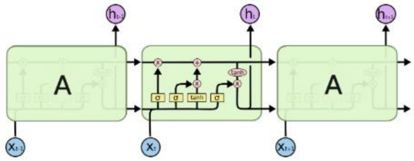

Fig. 3. Architecture of LSTM [9]

2.6 Convolutional Neural Network-CNN

CNN's were first developed and used around the 1980s. It is mainly used for image processing such as image recognition, image classification, object detection, face recognition etc. Figure 4 shows architecture of CNN.

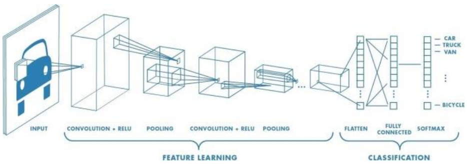

Fig. 4. Architecture of a CNN [10]

2.7 Support Vector Machine-SVM

Support vector machine is one of the widely used supervised machine learning algorithm. It is effective in high dimensional feature space, especially when the number of feature is greater than the number of training examples. The SVM tries to find a hyperplane that maximizes the margin between the support vectors. Different kernel functions can be used to train the SVM to handle non-linear input spaces. A simplified presentation of SVM is shown in Figure 5.

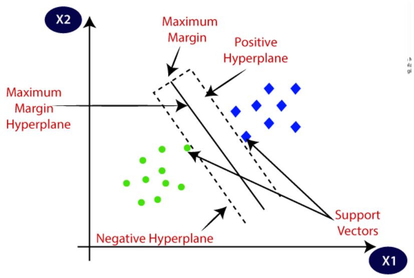

Fig. 5. Support Vector Machine [11]

2.8 Rough Set Theory

Rough set was first described by Pawlak [12]. It can be used to discover structural relationships in noisy data. RS is based on the establishment of equivalence classes within the given training data.

2.9 Moving Average Convergence/Divergence-MACD

MACD is a trend-following momentum indicator. It depicts the relationship between two moving averages of a variable. The Formula for MACD is:

MACD=12-Period EMA – 26-Period EMA

Figure 6 shows an example MACD Graph.

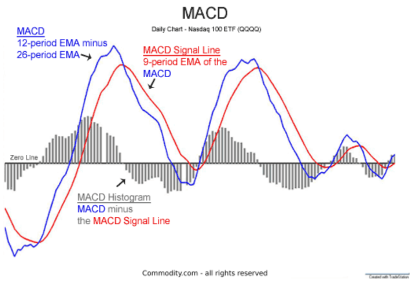

Fig. 6. MACD Graph [13]

2.10 Simple Moving Average-SMA

SMA is simply the average over a specified period. Popular SMAs are MA5, MA12, MA26 etc. An example of SMA is shown in the below Error! Reference source not found.

2.11 Rate of Change-ROC

The ROC calculates the difference between the current price and the price "n" periods ago. As the Rate-of-Change swings from positive to negative, the plot oscillates above and below the zero line. The formula of ROC is given below

$$ROC = [(Closet - Closet\text{-}n) / (Closet\text{-}n)] * 100$$

2.12 Relative Strength Index-RSI

Welles Wilder introduced RSI in June 1978. It is a type of momentum oscillator and is one of the most extensively used indicators in stock market technical analysis. The formula of Relative Strength is given below:

$$RSI = 100 - (100/1+RS)$$

RS (Relative Strength) = Average of X days up closes / Average of X days down closes

3 Literature Review

A deep learning model using LSTM was developed by Sachdeva et al. [14] for forecasting National Stock Exchange (NSE) of Indian equity market. The dataset contained data of NIFTY 50 index and INFOSYS Ltd from Yahoo finance ranging from 11 Dec 2007 to 12 Dec 2018. The features were Date, Open, High, Low and Close prices. To avoid snooping bias, they split the data into train and test sets. Train set contained data from 11 Dec 2007 to 11 Dec 2017, and test set had the remaining data. Min Max Scaler from sklearn library was used as sigmoid activation function is known to work well with scaled data. In the experiment two-time steps of 20 and 60 were used. The RNN model was fed ‘n’ stock prices before time‘t’ to analyze and capture correlation and trend in price movement. Based on its discovery the price of‘t+1’ day was predicted. Dropout was used for regularization and ‘Adam’ & ‘RMSProp’ were used as optimizers. For different combination of optimizer, time steps and LSTM, the best output was found for RMSProp with 60-time steps. They achieved 97% accuracy in predicting Infosys and Nifty 50 stock market results. Their test set Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) were 0.5625 and 0.74, respectively. From the results they reached at the conclusion that deep learning model is capable of predicting the NSE stock market.

Banik et al. [15] proposed a hybrid Rough Set & Artificial Neural Network model for Dhaka Stock Exchange (DSE) movement prediction and achieved 97.89% accuracy overall based on their dataset and experiment. They collected 8 years data of DSE spanning from January, 2004 to June, 2012. They spit the data to 70-30 train-test split. From the dataset, they generated MACD, MA5, MA12, PROC and RSI technical features which are generally used by technical analyst for financial time series prediction. Later these features were fed to train the models. To perform reducts and construct decision rules, the Rosetta Rough Set Toolkit was utilized to create a rough set model. 12 reducts were found from their data. The RS model provides 87% overall prediction accuracy with 72.31% of falling stock prices and 80.80% of rising stock prices prediction accuracy. For ANN they used 9:15:18:1 architecture where the learning rate used was 0.01 and momentum factor 0.90. They selected sigmoid transfer function at input and output layer. For back propagation, Levenberg Marquardt back propagation algorithm was chosen. The ANN model predicted declining stock prices with 72.86% accuracy, increasing stock prices with 80.74% accuracy, and overall stock prices with 78.37% accuracy. For the hybrid RS-ANN model first the ANN model was trained similar to previous architecture but with momentum factor 0.95. After that, the RS model was utilized to make reductions and construct decision rules based on the projected index. The hybrid model out performed both RS and ANN models with accuracy of 97.89%, 95.93% and 96.87% for overall, falling and raising stock price prediction respectively. Based on their experiment results they recommended the hybrid model as a tool to guide investors of DSE.

In their attempt to predict the Karachi Stock Exchange (KSE), Usmani et al. [16] used Single Layer Perceptron (SLP), Multi-Layer Perceptron (MLP), Radial Basis Function (RBF) and Support Vector Machine (SVM). They use 9 parameters i.e., Oil rate, gold rate, silver rate, Interest rate, Foreign Exchange Rate (FEX), News, Social Media feed, Simple Moving Average (SMA), Autoregressive Integrated Moving Average (ARIMA). They found that Oil rate has highest co-variance with market trend whereas FEX showed co-variance neutral. The attribute values were normalized between -1 to 1 as all of them can have negative/positive values. The dataset span from September 2015 to January 2016, because of dependencies on two attributes i.e., News and Social Media Feed (twitter), they were able to collect only 100 records of them between the mentioned time. These data were split to 70-30 ratio for train-test sets. Opinion Finder library was used for the text mining which lie between day close to next day market opening. As there were many non-English wards in the corpus, hey suspected it might have compromised the models' performance. Among the four developed models, MLP has the highest 77% accuracy on the test set. But SVM surprisingly showed 100% accuracy on the training set even after when the library was changed. However, its performance on test set was lowest 60%.

In their paper Bhowmick et al. [17] explored the prediction efficiencies between classical statistical models for financial forecasting and advanced neural networks namely RNN, LSTM, LSTM peephole and Gated recurrent unit (GRU). They came to a conclusion that NN models outperform classical models such as ARIMA. Which shows the nonlinearity of the market was better captured by the NN models. The compare RMSE for performance evaluation, among the four NN models, performance was close. The dataset comprised of 5 companies from DSE. The NAN values were dropped before applying MinMaxScaler from sklearn library. Four prices, Average, Close, Low & High were used as input to the models. The split into 80-10-10% for train-test-validation set, they used a 4-20-50-1 architecture for the NN models with learning rate 0.00001. Adam optimizer was used as it is less memory intensive and much suited for non-stationary, sparse data. The experiment was run for 100 epochs. For NN models they used tensorflow and keras. Statsmodels library was used to train classical models such as autocorrelation (acf), partial-autocorrelation (pcf), the first differential autocorrelation and ARIMA. To further validate the experiment, they used Moving Average Convergence/Divergence (MACD) indicator, but later it was considered uninformative for prediction.

Selvin et al. [18] proposed a model for short term value prediction of the NSE listed companies. They developed models for RNN, LSTM and CNN with sliding window approach and found CNN model was capable of identifying interrelation within the data and predict stock values. The dataset comprised of minute wise stock data of 1721 NSE listed companies between the period of July'14 and June'15. However, the work focused on 3 companies, 2 from the IT sector and 1 from the pharmaceuticals sector. The sliding window size was chosen to be 100 minutes which was chosen based on error rate of various window sizes. Interestingly, the training data was of Infosys from the period of 1 July 2014 to 14 October 2014. But the test data consisted of all three companies chosen earlier and lied between 16 October 2014 and 28 November 2014. All models were trained for 1000 epochs. From the results it was evident that CNN gave more accurate results than other two models. It was asserted that the superior accuracy of CNN was due to its ability to predict the value based on current window. This enabled CNN to capture the dynamic changes and patterns occurring in current window. The NN models' result was compared with the output of ARIMA and it was found NN models outperformed ARIMA.

In their work, Khare et al. [19] built models to forecast short-term prices of stocks in New York Stock Exchange using two distinct sorts of ANNs, Feed Forward Neural Networks and RNNs. The dataset consisted of over a one-year period long, minute-by-minute stock price data for 10 unique stocks recorded on the New York Stock Exchange. Normalization was done using MinMaxScaler to fit the data between 0 and 1 before feeding it to the NN models. Though they did not mention which particular features were fed to the models, the mentioned three categories Trend, Oscillator and Momentum, to which the features belong. The dataset was split to 70-30 ratio for train-test set. From the results of the experiment, they found that both LSTM and MLP models were able to predict the up-down price trend. However, the MLP model was able to predict the price more closely compared to the LSTM model.

Li et al. [20] used 6 forecasting models with 9 feature combinations selected from 23 commonly used stock market technical indicators to analyze China Stock Market. They employed 12 distinct time periods to assess the impact of time period on the forecasting methods, using 400 epochs on a moving-window basis. Additionally, they compared the relationship between different feature combinations and model performance. In their work, they have tried to compare the performance of shallow machine learning (SVM, Naive Bayes and Decision Tree) and deep learning (MLP, RNN, LSTM). Deep learning modes outperformed other models in terms of accuracy, according to their findings.

Porshnev et al. [21] tested the hypothesis that twitter sentiment analysis could provide additional information to increase stock market prediction accuracy. Using SVM the achieved the best accuracy of 64.10%, therefore rejecting their earlier hypothesis. They pointed the lack of accuracy can be attributed to low data points.

4 Conclusion

In this paper we have tried to summarize several concepts explored by researchers in the fields. We have showed how different stock exchange data are analyzed and which algorithms perform better. Most of the research is focused on a single company's data. However, most of the time there is a relationship between the price movements of more than one company. This is something that the researchers can explore.

References

[1] Hayes, A. (2021). Stock Definition. *https://www.investopedia.com/terms/s/stock.asp*.
[2] Chen, J. (2022). Stock Market Definition. *https://www.investopedia.com/terms/s/stockmarket.asp*.
[3] Stock Market - What is the Stock Market and how it Works. *https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/stock-market*.
[4] Machine learning - Wikipedia. *https://en.wikipedia.org/wiki/Machine_learning*.
[5] Artificial Neural Network - Applications, Algorithms and Examples - TechVidvan. *https://techvidvan.com/tutorials/artificial-neural-network*.
[6] Recurrent Neural Networks | IBM. *https://www.ibm.com/cloud/learn/recurrent-neural-networks*.

[7] File:Recurrent neural network unfold.svg - Wikimedia Commons. *https://commons.wikimedia.org/wiki/File:Recurrent_neural_network_unfold.svg*.

[8] Hochreiter, S. and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*,9(8):1735–1780.

[9] Neural networks - Difference between feedback RNN and LSTM/GRU - Cross Validated. *https://stats.stackexchange.com/questions/222584/difference-between-feedback-rnn-and-lstm-gru*.

[10] Understanding of Convolutional Neural Network (CNN) — Deep Learning | by Prabhu | Medium. *https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148*.

[11] Support Vector Machine (SVM) Algorithm - Javatpoint. *https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm*.

[12] Pawlak, Z. (1982). Rough sets. *International Journal of Computer & Information Sciences*, 11(5):341–356.

[13] When to Use and How to Read the MACD Indicator - Commodity.com. *https://commodity.com/technical-analysis/macd/*.

[14] Sachdeva, A. et al. (2019). An Effective Time Series Analysis for Equity Market Prediction Using Deep Learning Model. In *International Conference on Data Science and Communication*, 1–5.

[15] Khan, K., Anwer, M. and Banik, S. (2013). Market Timing Decisions by Hybrid Machine Learning Technique : A Case Study for Dhaka Stock Market.

[16] Usmani, M. et al. (2016). Stock Market Prediction Using Machine Learning Techniques. In *3rd International Conference on Computer and Information Sciences,* 322-327,

[17] Bhowmick, A., Rahman, A. and Rahman, R. M. (2019). Performance analysis of different recurrent neural network architectures and classical statistical model for financial forecasting: A case study on dhaka stock exchange. In *Artificial Intelligence Methods in Intelligent Algorithms,* 985. Springer International Publishing.

[18] Selvin, S. et al. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. in International Conference on Advances in Computing, Communications and Informatics*,* 1643–1647.

[19] Khare, K. et al. (2017). Short Term Stock Price Prediction Using Deep Learning. In *2nd International Conference on Recent Trends in Electronics Information & Communication Technology,* 482–486.

[20] Li, W. and Liao, J. (2017). A comparative study on trend forecasting approach for stock price time series. In *Proceedings of the International Conference on Anti-Counterfeiting, Security and Identification,* 74-78.

[21] Porshnev, A., Redkin, I. and Shevchenko, A. (2013). Machine learning in prediction of stock market indicators based on historical data and data from twitter sentiment analysis. In *Proceedings - IEEE 13th International Conference on Data Mining Workshops,* 440–444.