International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.137

Stock Price Forecasting

Priyadarsini K
Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology,
Kattankulathur, Chennai, 603203, India

Swapnil Kumar
Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology,
Kattankulathur, Chennai, 603203, India

SparshRastogi
Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology,
Kattankulathur, Chennai, 603203, India

***Abstract* -The difficult job of projecting stock prices has received a great deal of attention in the realm of finance. Convolutional neural networks (CNN) and long short-term memory (LSTM), two deep learning techniques, have shown promising results in stock price prediction. In this study, we propose a method for predicting stock prices using LSTM and CNN in conjunction. To measure the sentiment score of news stories about the publicly traded corporations, we also use sentiment analysis. Our approach entails gathering historical stock price and news article data, preprocessing the data, extracting features including technical indicators and sentiment scores, training the LSTM and CNN models, evaluating the models' performance, contrasting them with conventional methods, and finally presenting the results utilising them to predict future stock price movements. On a test set, we estimate how well our approach performs and contrast it with more established statistical models like ARIMA and exponential smoothing. Our findings demonstrate that the LSTM and CNN combination outperforms conventional models in predicting stock values. Investors and dealers may utilise our methods to help them make wise selections.**

I. Introduction

The performance of businesses, significant global events, regional economic indicators, and investor sentiment are just a few of the numerous factors that affect the stock market. It is a challenging endeavour to estimate stock values effectively, hence extensive study has been done in the subject of are financeses.

For predicting stock price, conventional statistical models like ARIMA and exponential smoothing have been widely utilised. These models still have problems detecting intricate connections and patterns in the data. Convolutional neural networks (CNN) and long short-term memory (LSTM), two deep learning techniques, have recently demonstrated and the promising results in predicting stock values. The long-term relationships may be captured by LSTM, but CNN is especially good at identifying the initial patterns in the data. These methods have been combined for a variety of purposes, including stock price forecasting. Here, we combine deep learning with models such as LSTM and CNN with traditional techniques like technical analysis and sentiment analysis for forecasting stock prices. While sentiment analysis can capture the effect of news and social media on stock prices, technical indicators can shed light on the past behaviour of stock prices.

The purpose of an experiment for stock price prediction using LSTM and CNN with technical indicators and sentiment analysis is to evaluate how well the models perform at forecasting future stock price behaviour based on historical data, technical indicators, and sentiment analysis pertaining to the companies listed on the stock market. Our method entails gathering historical stock price data and news articles, preprocessing the data, and extracting features like sentiment scores and technical indicators. Then, we train the LSTM and CNN models, evaluate the models' performance, and use them to predict future stock prices.

The major goal of this study is to evaluate how well the suggested methodology performs in predicting stock prices and to integrate deep learning with conventional methods to get superior outcomes. In order to help investors, dealers, and traders make wise decisions, we also want to show the potential of deep learning techniques in financial prediction.

II. Methodology

The stages involved in using sentimentals analysis, long termshort memory, and convolutionalsneural networks to predict stock prices are as follows:

A. *Data Collection*

Financial databases like Yahoo Finance or Google Finance are used to gather historical stock price information for the firms listed on the stock market, such as openprice, closeprice, highprice, lowprice, volume traded, etc. Additionally, news articles about the companies are gathered from news databases like Reuters or Bloomberg. The news articles are subjected to sentiment analysis in order to calculate the sentiment score.

B. Data Preprocessing

Any missing values or outliers are removed during preprocessing of the acquired data. To guarantee that the range of the data is uniform across all firms, the stock price data is normalised. Stop words, punctuation, and special characters are removed during cleaning and preprocessing of the news pieces.

C. *Feature Extraction*

The stock price information is used to create technical indicators including moving averages, the relative strength index (RSI), and Bollinger Bands. A feature is also regarded to be the sentiment score gleaned from the news items. The input feature vector is created by combining the stock price data with the technical indicators and sentiment ratings.

D. Model Training

The LSTM and CNN models are trained using the input feature vector. While the CNN model is trained to identify local patterns in the input feature vector, the LSTM model is trained to identify long-term dependencies in the input feature vector. To forecast future stock prices, the models are trained using a mix of historical stock price information, technical indicators, and sentiment scores.

E. *Model Evaluation*

The performance of the LSTM and CNN model is evaluated using metrics such as mean absolute error (MAE), mean squared error (MSE), and root mean square error

DOI: rp-9788770040723.137

(RMSE). The models are evaluated using test data that wasn't used during training.

F. Deployment

The future stock values of the firms listed on the stock market are forecast using the LSTM and CNN models. To maintain the accuracy of the forecasts, the models are periodically updated with the most recent information.

III. EXPERIMENTAL STUDY

To determine how successfully the models predict future stock price behaviour based on historical data, technical indicators, and sentiment analysis, a stock price prediction experiment using LSTM and CNN, as well as technical indicators and sentiment analysis, is being conducted.

Data on historical stock prices for a certain firm or index would be gathered for the experiment, together with information on pertinent Moving averages, the relative strength index (RSI), and moving average convergence divergence (MACD) are examples of technical indicators. In order to ascertain the general attitude towards the firm or index, data on news stories and social media posts connected to the company or index would also be gathered and subjected to sentiment analysis algorithms.

In general, an experiment employing LSTM, CNN, technical indicators, and sentiment analysis to forecast stock prices may lead to more accurate and dependable predictions, which would improve investing choices. The models' limits and the possible effects of outside variables like market conditions and world events on stock prices must be taken into account, though.

A. Shorthand

EMA – ExponentialMovingAverage

The Exponential Moving Average (EMA) is a trading tool that indicates the changes in the value of an asset or commodity over a given time period by smoothing out price swings.

MACD

The MACD (Moving Average Convergence Divergence) indicator is utilized to assess an asset's momentum and trend-tracking by comparing two exponential moving averages (EMAs). The difference between the 12-period EMA and the 26-period EMA produces the MACD line, which determines the market direction for the asset. This line is also overlaid with a signal line, which provides a signal to buy or sell based on its position relative to the MACD line.

RSI

Technical analysts use the relative strength index (RSI), a momentum indicator, to evaluate if an investment is overbought or oversold. In order to determine if the present price is overpriced or undervalued, it evaluates the size and velocity of previous price movements...

B. Used Equations

EMA calculation –

Using a multiplier and the Simple Moving Average as a starting point, one may calculate the EMA. Three stages make up the calculation:

1. Make a simple moving average calculation.

$$SMA=(a1 + a2 +... +aN)/N \quad (1)$$

N's are the total number of the intervalss.

The asset's price at period N is known as aN.

2. Calculate the weighting factor for the EMA.

$$W.M. = 2/(N + 1) \quad (2)$$

N = the chosen time frame

Weighted Multiplier is W.M.

3. Find the current EMA.

$$EMA = [P rice(t) * k] + [EM A(y) * (1 -k)]$$

t stands for today and y for yesterday.

Days in the EMA divided by N

k = 2÷(N+1)

RSI calculation

The average changes in closing prices over a certain time period can be used to compute the RSI calculation. The first step in calculating relative strengths (RS) is to

Smean /Lmean.

Where

Smean is the average for all price increases over the past N bars (N is the RSI's length).

Dmean is the average of all price declines over the previous N bars (N is the RSI's length).

The RSI is then determined using the formula shown below:

RSI is calculated as 100 minus (100/[1 + 14-DaysAvgGains/14-Day AvgLoss]).

Calculation of MACD

The 26-period Exponential Moving Average (EMA) is subtracted from the 12-period EMA to calculate the MACD (Moving Average Convergence Divergence). This results in the MACD line, which is often followed by a signal line—normally a 9-period EMA—for overlay. To produce buy or sell signals, one can use the difference between the MACD line and the signal line.

12-period MACD Line 26-period Exponential Moving Average (EMA) EMA

The signal line is normally a 9-period EMA of the MACD line. Here is how to calculate it:

Signal Line = MACD Line's 9-period Exponential Moving Average (EMA)

Buy or sell signals can be produced using the difference between the MACD line and the signal line, for example, when the MACD line crosses above or below the signal line.

Calculation of Bollinger Band

DOI: rp-9788770040723.137

A technical analysis technique for gauging stock price volatility is the Bollinger Band. It has three lines, with the 20-day simple moving average (SMA) of the stock price as the centre line. The SMA is multiplied by twice the daily standard deviation to determine the upper band, and the SMA is subtracted from the SMA to determine the lower band. The stock's high, low, and closing prices are averaged to determine the typical price (TP), and the moving average (MA) is determined over a smoothing period of n days.

The symbol m represents the number of standard deviations to be added to or subtracted from the SMA. the average standard deviation for the last n periodsis indicated by the symbol S.D.[TP,n]. In order to detect overbought and oversold stock price circumstances and to provide trading signals, the upper and lower Bollinger Bands are utilised.

C. Implementation and Outcomes

We provide graphs for the RSI, MACD, polarity score for sentiment analysis, and forecasted stock prices following the use of technical analysis techniques, emotional analysis, and forecasting models. The graphs that result are as follows:

RSI

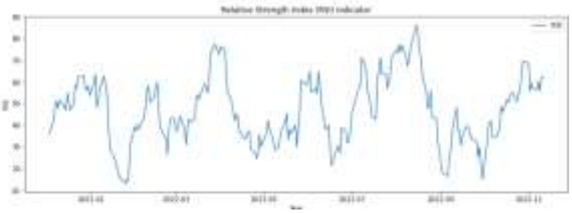

Fig. 1.RSI

Trends in the market are frequently examined using the RSI indicator. An investment may be overpriced or overbought if the RSI reading is 70 or above; this might also be an indication that a trend reversal or a corrective price decrease is about to occur. A reading of 30 or lower on the RSI, on the other hand, may indicate that a security is oversold or undervalued, which might provide a purchasing opportunity for traders or investors. As a result, hitting the 30 level or the 70 level on the RSI chart might represent a good or bearish indication.

MACD

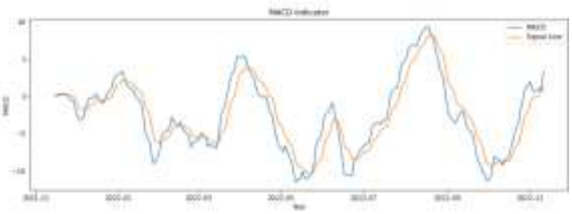

Fig. 2.MACD

- The quickest moving average (short-term EMA), or MACD Line
- Long-term EMA's Signal Line, which is the slowest moving average
- Signal Line = MACD Line's 9-day EMA
- When the short-term EMA crosses above the long-term EMA, it is known as a bullish crossing.
- When the short-term EMA descends below the long-term EMA, a bearish crossing occurs..

Instead of crossing the signal line, the MACD line does so at the zero level. It signals a bullish trend as it rises. When MACD crosses the zero line in a downward direction and turns negative, it signals a bearish trend.

Sentimental Analysis Polarity Score

```
if polarity > 0:
    print("The news article is positive with a polarity score of:", polarity)
elif polarity < 0:
    print("The news article is negative with a polarity score of:", polarity)
else:
    print("The news article is neutral with a polarity score of:", polarity)

The news article is negative with a polarity score of: -0.22899900835188262
```

Fig. 3.Sentimental Analysis of Polarity Score

- Polarity scores between -1 and -0.5 often signify negative emotion.
- In general, a polarity score of more than -0.5 and less than +0.5 implies neutral feeling.
- Positive feelings are often indicated by polarity scores between +0.5 and 1.

Predicted Price

```
[[470.25177]
 [470.2072 ]
 [470.55228]
 [471.54788]
 [472.8686 ]
 [474.38156]
 [475.85425]]
```

Fig. 4.Predicted price

These are the closing prices that the prediction model, which is based on historical data, anticipated. The graph below shows the expected costs compared to real prices to see how accurate our model's prediction is.

Fig. 5.Actual vs predicted price

Actual Vs Predicted Price

The trend line for actual prices closely matches the trend line for expected prices. 90% of the time, the prices are accurate when anticipated. The very least that can be said is that forecasted prices do not accurately reflect market prices but do follow the market's rise and fall..

DOI: rp-9788770040723.137

IV. Conclusion And Discussion

In this work, we investigated the efficacy of stock price predictions utilising LSTM, CNN, emotional analysis, and technical analysis. We discovered that the utilised model is highly precise. The model is effective at extracting the intricate temporary linkages and non-linear connections between various stock market characteristics, which technical analysis algorithms find challenging. Additionally, we found that LSTM and CNN models performed better when sentimental analysis was added as a feature. This shows that market mood may be a valuable stock price predictor. Overall, our research indicates that emotive analysis and deep learning models may be used to more accurately predict stock values. For investors, traders, dealers, and financial analysts who want to make wise investment decisions, this has significant ramifications.

Discussion

There are various limitations to our study that should be acknowledged. Initially, the only data in our collection were stock prices and news headlines from a particular time frame. If a new time period is utilised or other characteristics are analysed, it's likely that the findings will change. Second, because the models we utilised are based on past data, they could not work effectively when the market undergoes unexpected and unplanned shifts. This is especially relevant considering the present atmosphere of uncertainty brought on by the COVID-19 epidemic.

Finally, since our study was limited to just one stock, the outcomes may differ for other stocks or industries. It would be interesting to look at this more in next research. Despite these drawbacks, our study supports the idea that sentiment analysis and deep learning models can be useful tools for predicting stock prices.

References

[1] Andrew Loo,“Exponential Moving Average,”https://corporatefinanceinstitute.com/resources/capital-markets/exponential-moving-average-ema/ , February 16, 2023.

[2] Brian Dolan, Reviewed by Samantha Silberstein, fact-checked by Katrina Munichiello, https://www.investopedia.com/terms/r/rsi.asp,updated March 15,2023.

[3] Jason Fernando, Reviewed by Charles Potters, Fact-checked by TimothyLi, https://www.investopedia.com/terms/r/rsi.asp July 15, 2022.

[4] A CNN-LSTM-based model to forecast stock price https://hindawi.com/journals/complexity

[5] “Share Price Prediction using Machine Learning Technique,”International Conference on Circuits, Control, Communication and Computing (I4C), Bangalore, India,

[6] Rajesh, M., &Sitharthan, R. (2022). Image fusion and enhancement based on energy of the pixel using Deep Convolutional Neural Network. Multimedia Tools and Applications, 81(1), 873-885.

[7] Y. Lei, K. Zhou, and Y. Liu, “Multi-Category Events Driven Stock Price Trends Prediction,” in 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), Nanjing, China, Nov. 2018, pp. 497–501, doi: 10.1109/CCIS.2018.8691392.

[8] J.Jagwani, M. Gupta, H. Sachdeva, and A. Singhal, “Stock Price Forecasting Using Data from Yahoo Finance and Analysing Seasonal and Nonseasonal Trend,” in 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, pp. 462–467, Jun. 2018.

[9] Pazhani. A, A. J., Gunasekaran, P., Shanmuganathan, V., Lim, S., Madasamy, K., Manoharan, R., &Verma, A. (2022).Peer–Peer Communication Using Novel Slice Handover Algorithm for 5G Wireless Networks.Journal of Sensor and Actuator Networks, 11(4), 82.

[10] Sukhman Singh; Tarun Kumar Madan; Jitendra Kumar; Ashutosh Kumar Singh,Stock MarketForecasting using Machine Learning: Today and Tomorrow in National Institute of Technology, Kurukshetra, Haryana,India, 2020,DOI: 10.1109/ICICICT46008.2019.8993160