International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Stock Market Prediction Using LSTM

Rajnish Sharma, Sakshi Jain, Sanskriti Singh, Nitin Kumar, Megha

Department of Computer Science and Engineering, Quantum School of Technology, Quantum University Roorkee, India

ABSTRACT

This research explores the application of Long Short-Term Memory (LSTM) neural networks for predicting stock market prices based on historical data. RNN is a modern technology that is use in it, traditional models often fail to capture long memory capacity. LSTM, a type of recurrent neural network, is specifically designed to handle long-term dependencies in time-series data. In this study, historical stock data such as opening price, closing price, high, low, and trading volume were used to train the LSTM model. The results indicate that LSTM networks provide superior accuracy and robustness in capturing the sequential behaviour of stock prices over time, and the results indicate that LSTM can effectively predict Long-term trends in stock prices. This approach demonstrates potential for use in financial forecasting and decision-making systems.

Keywords— LSTM, Stock Market Prediction

Introduction

Stock prediction is a fundamental concept that use in real life predictions behalf on the historical data. The stock market plays a critical role in the global economy, influencing investment decisions, marketing decisions and mainly focus on the economic growth, and wealth distribution. **Long Short-Term Memory (LSTM)** networks are a class of deep learning models that excel at learning from sequential data, especially when long-term dependencies exist. Unlike traditional models, LSTM does not require the input time series to be stationary and can handle complex, non-linear relationships without the need for heavy manual feature engineering. LSTMs are specifically designed to remember information over long sequences. In stock markets, past price trends and patterns can influence future movements, even after many time steps. LSTM's memory cells and gating mechanisms enable it to retain relevant past information more effectively than traditional RNNs or time series models. LSTM is a very accurate and good method of stock prediction in place of rest terminology, which was used in our economic growth analysis, Future Benefits in Stock Marketing. Traditional time series models often require the data to be stationary (i.e., stable mean and variance over time), which is rarely the case in stock markets. LSTMs can handle nonstationary inputs without the need for extensive pre-processing or transformation. LSTM method is a Beneficial for the Live Marketing Techniques.it is a great way to enhance business rules by the training and testing to data parameters for long term predictions. Most secure for both (Long and Short) memory predictions.

With online learning techniques and frequent retraining, LSTMs can be adapted for Realtime or near-real-time stock price prediction, making them suitable for applications in algorithmic trading and live market analysis.

Long Short-Term Memory (LSTM):

LSTM (Long Short-Term Memory) is a specialized type of RNN introduced to solve the problems of standard RNNs, particularly the **difficulty in learning long-term dependencies**.

LSTMs use a more complex architecture that includes **memory cells** and **gating mechanisms** to control the flow of information.

Firstly, we need to collect data for the analysis the results based on financial, Marketing, and business Understandings.

Key Components of LSTM:

1. **Forget Gate**: Controls which information from the previous state is retained for the current state. A sigmoid function determines which information will be forgotten and which will be retained.
2. **Input Gate:** Decides what new information should be stored in the cell state.
3. **Cell State:** The internal memory that carries long-term data.
4. **Output Gate:** Determines what part of the cell state should be output.

An **LSTM** is more like long-term memory: it can remember important events The core idea behind LSTM is its **cell state**, which acts like a long-term memory. It can carry information across many time steps without forgotten quickly.

Traditional machine learning models like Linear Regression, ARIMA, SVM, Random Forest, and K-Nearest Neighbours (KNN) are generally not suitable for long-term stock prediction due to several limitations. Linear Regression Sometime these can be beneficial for data analysis process but for the long-term analysis these may be fail in process, assumes a linear relationship between input and output, which oversimplifies the highly nonlinear and volatile nature of stock markets. It cannot capture sudden market shifts or long-term trends effectively. ARIMA, while useful for short-term forecasting, relies on the assumption that the time series is stationery and univariate. It fails to incorporate external factors such as economic indicators or news sentiment, which are crucial for long-term forecasting. Additionally, ARIMA struggles with capturing nonlinear and complex market dynamics.

Support Vector Machines (SVM) are better suited for classification problems and do not naturally handle sequential or time dependent data, which is essential for stock prediction. Moreover, as the amount of data increases, SVM becomes computationally expensive and hard to tune for optimal performance. Random Forests are powerful for many classification and regression tasks but treat each data point independently. This means they ignore the temporal structure of stock data, leading to poor performance over extended time horizons. They are also prone to overfitting in noisy and volatile datasets, which is a common characteristic of long-term stock data, But after a long time it can be shows buffer in analytical process due to the some Long term terminologies.

K-Nearest Neighbour (KNN), on the other hand, is a very simplistic algorithm that works based on similarity with past data points. It has no concept of time or sequence and becomes highly inefficient and inaccurate when dealing with largescale, long-term data. Since KNN doesn't learn a model but rather memorizes data, it doesn't generalize well to future trends compare to LSTM terminologies.

In contrast, models like Some LSTM, GRU, or Transformer based architectures are more suitable for long-term stock prediction. These models can capture complex, nonlinear patterns and long-term dependencies in sequential data. They are capable of learning from multiple dimensional Data phases adapting to changes over time, and providing more robust predictions in volatile markets that effective for the Long-term prediction and these can contain too old dataset to analysis the data that depends over time and financial path.

Methodology and Data

The methodology for predicting stock prices using an LSTM neural network involved several key stages: data collection, pre-processing, model development, training, and evaluation. The historical stock data was sourced from Yahoo Finance using the Python library. The dataset included essential daily past trading information such as the opening price, highest price, lowest price, closing price, adjusted closing price, and trading volume of the price data.

To build our model we are going to use the LSTM RNN, our model uses 70% of data for training and the other 30% of data for testing. our data series cover the period going from **4/01/2010** to **6/12/2024**. Data covers almost 15 Years old data.

Our model will be structured as follow:

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 100, 50)	10,400
dropout (Dropout)	(None, 100, 50)	0
lstm_1 (LSTM)	(None, 100, 60)	26,640
dropout_1 (Dropout)	(None, 100, 60)	0
lstm_2 (LSTM)	(None, 100, 80)	45,120
dropout_2 (Dropout)	(None, 100, 80)	0
lstm_3 (LSTM)	(None, 120)	96,480
dropout_3 (Dropout)	(None, 120)	0
dense (Dense)	(None, 1)	121

Figure 1: the LSTM model structure

The model architecture consists of four LSTM layers with varying units (50, 60, 80, and 120) and dropout regularization (ranging from 0.2 to 0.5) to prevent overfitting. The final layer is a dense layer with one unit for the output. The model's summary, detailing the number of parameters in each layer, indicates total of 178,761 parameters.

Result and discussion:

The implementation of proposed LSTM model using python which predicts the financial conditions of YAHOO share based on its historical data. The below visualization figure shows the visualization of stock prediction. In our paper the implementation of an algorithm which predicts the stock price of a share for given period of time, the below graph from our algorithm will show the predicted price of TCS (YAHOO Finance) share. In the result shown in the below graph is the plotted form our algorithm outcome by applying LSTM for achieving the accuracy. the “x” axis is Share time. The “y” axis is Price. The data is slot of 15 years.

The graph successfully along with the predicted price original price (blue) and Predicted price (red).

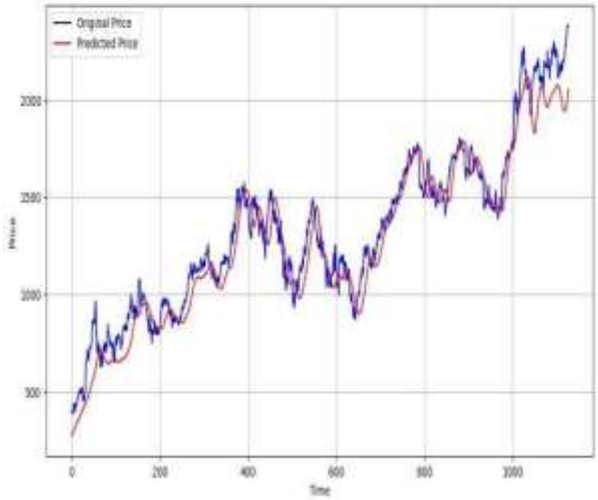


Figure 3: result of training for the NKE stocks with different dataset time

Compared to traditional models such as ARIMA and linear regression, the LSTM model consistently achieved lower error values, indicating better predictive accuracy. The visualization of predicted prices against actual prices further supported these results, showing that the LSTM model was able to track movements in the stock market with reasonable precision. Although the model did not perfectly predict sudden spikes or drops—common limitations due to market randomness—it showed significant improvement over statistical models in overall trend prediction. These results highlight the suitability of LSTM for time-series forecasting in highly volatile and non-linear domains like the stock market.

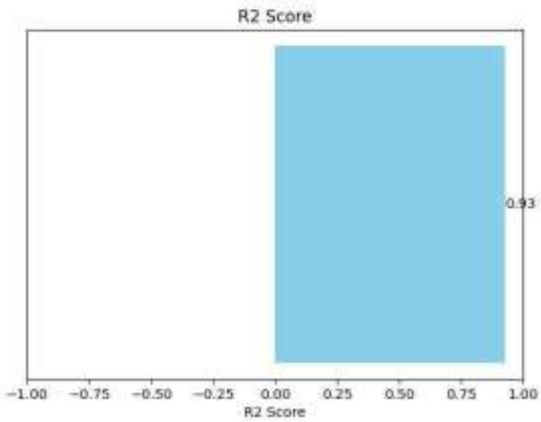


Figure: Accuracy chart of LSTM

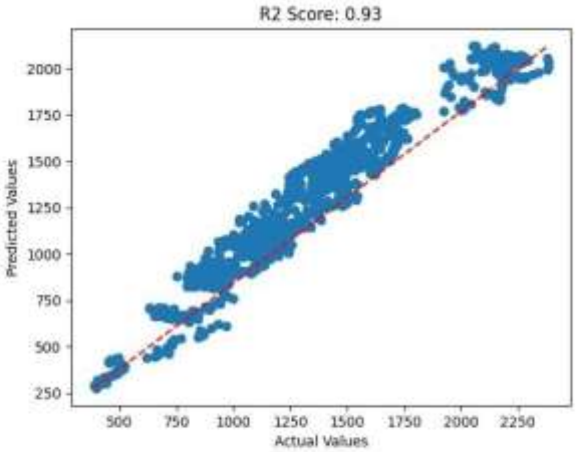
The financial market is a highly dynamic and complex system influenced by numerous factors, including economic indicators, market sentiment, and global events. Predicting stock prices has long been a challenging task for researchers, investors, and data scientists. With the advent of machine learning, significant progress has been made in stock market behaviour through data driven approaches.

This research focuses on the use of supervised regression techniques for predicting stock prices based on historical data. To know the predict values and the actual values we apply the Scatter plot Graph, The goal is to build a model capable of capturing the underlying trends in the data to forecast future prices with high accuracy. Regression algorithms are particularly suited for such tasks as they attempt to find the best-fit mapping between the input features and the target price.

In this study, historical stock data was used to train and evaluate a regression-based predictive model. The performance of the model was assessed using the coefficient of determination (R^2 score), which reflects how well the model's predictions align with the actual values. The model achieved an R^2 **score of 0.93**, indicating a high degree of accuracy and suggesting that it successfully captured key patterns within the data.

This work highlights the potential of machine learning financial forecasting and demonstrates how regression models can be effectively applied to stock price prediction. The promising results lay a strong foundation for further exploration, including more advanced algorithms, feature engineering, and real-time deployment in financial decision-making systems. A scatter plot Graph define about the accurate predict and actual values.

The scatter plot Graph given as follow:



Conclusion:

Long term stock trading prediction is a non-trivial task with way less attention than it should be. This paper proposes a model with LSTM and fully connected layers to predict long term stock predictions based on financial statements. Experimental results show that LSTM technique is powerful in training models on structured data, erasing partial information from training examples. For future work, more historical data in financial statements could be critical in terms of improving the model's accuracy, like extending 3 data points in an item to 10 data points in an item for model training. Predicting stock performance in an even longer period of time, like 2 years, 5 years, or 10 years, even 1lmost 15years that we apply into LSTM Method

References:

1. Batres-Estrada, B. (2015). Deep learning for multivariate financial time series.
2. Emerson, S., Kennedy, R., O'Shea, L., & O'Brien, J. (2019, May). Trends and Applications of Machine Learning in Quantitative Finance. In 8th International Conference on Economics and Finance Research (ICEFR 2019).
3. Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33(1), 3-12.
4. Moritz, B., & Zimmermann, T. (2016). Treebased conditional portfolio sorts: The relation between past and future stock returns. Available at SSRN 2740751.
5. Olah, C. (2015). Understanding lstm networks–colah's blog. Colah. github. io.
6. Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2018). Decision- Making for Financial Trading: A Fusion Approach of Machine Learning and Portfolio Selection. Expert Systems with Applications.
7. Patterson J., 2017. Deep Learning: A Practitioner's Approach, O'Reilly Media.
8. Siami-Namini, S., & Namin, A. S. (2018). Forecasting economics and financial time series: Arima vs. lstm. arXiv preprint arXiv:1803.06386.
9. Takeuchi, L., & Lee, Y. Y. A. (2013). Applying deep learning to enhance momentum trading strategies in stocks. In Technical Report. Stanford University.
10. Wang, S., and Y. Luo. 2012. "Signal Processing: The Rise of the Machines." Deutsche Bank Quantitative Strategy (5 June).
11. Ryo Akita, Akira Yoshihara, Takashi Matsubara, and Kuniaki Uehara. Deep learning for stock prediction using numerical and textual information. In 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), pages 1–6. IEEE, 2016.
12. Kai Chen, Yi Zhou, and Fangyan Dai. A lstm-based method for stock returns prediction: A case study of china stock market. In 2015 IEEE international conference on big data (big data), pages 2823– 2824. IEEE, 2015.

13. Xiao Ding, Yue Zhang, Ting Liu, and Junwen Duan. Deep learning for event-driven stock prediction. In Twenty-fourth international joint conference on artificial intelligence, 2015.

14. Robert W Holthausen and David F Larcker. The prediction of stock returns using financial statement information. Journal of accounting and economics, 15(23):373–411, 1992.

15. Jane A Ou and Stephen H Penman. Financial statement analysis and the prediction of stock returns. Journal of accounting and economics, 11(4):295–329, 1989.

16. Eberhard Schöneburg. Stock price prediction using neural networks: A project report. Neurocomputing, 2(1):17–27, 1990. Robert P Schumaker and Hsinchun Chen. A quantitative stock prediction system based on financial news.