IBM Stock Price Prediction Using Long Short-Term Memory (LSTM) Networks

Author: Gursimran Sujlana

Abstract

Stock price prediction has always been a challenging task due to the volatile nature of the stock market. This research explores the use of Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Network (RNN), to predict the stock prices of IBM. The model is trained on historical stock price data and aims to provide accurate future stock price predictions. The results demonstrate the potential of LSTM networks in capturing temporal dependencies in time series data for stock price forecasting.

1. Introduction

The stock market is inherently complex and influenced by numerous factors, making accurate prediction of stock prices a formidable task. Traditional statistical methods often fall short in capturing the non-linear relationships and temporal dependencies inherent in stock data. Recent advancements in deep learning, particularly Long Short-Term Memory (LSTM) networks, have shown promise in modeling such dependencies and improving prediction accuracy.

2. Literature Review

Previous studies have explored various machine learning techniques for stock price prediction, including Support Vector Machines (SVM), Random Forests, and Neural Networks. However, these methods often struggle with sequential data. LSTM networks, designed to handle long-term dependencies, have been applied successfully in fields such as natural language processing and speech recognition, and are now being explored for financial time series forecasting.

3. Methodology

3.1 Data Collection and Preprocessing

The dataset used in this study consists of historical stock prices of IBM, split into training and testing sets. The training data spans from the start date to a specific cutoff date, while the test data covers the period immediately following the training data.

3.2 Data Normalization

To prepare the data for the LSTM model, the stock prices were normalized using MinMaxScaler, scaling the values between 0 and 1. This step is crucial to ensure that the model performs optimally.

3.3 Creating Timesteps

A sliding window approach with 60 timesteps was used to create the input data for the model. This means that the model uses the past 60 days of stock prices to predict the next day's price.

3.4 Model Architecture

An LSTM network was constructed with the following layers:

- Four LSTM layers with 50 units each.
- Dropout layers with a dropout rate of 20% to prevent overfitting.
- A dense output layer with a single neuron to predict the stock price.

The model was compiled using the Adam optimizer and the mean squared error loss function. The training process involved fitting the model on the training data for 100 epochs with a batch size of 32.

3.5 Model Training and Evaluation

The model was trained on the training dataset, and its performance was evaluated on the test dataset. The predictions were compared with the actual stock prices to assess the model's accuracy.

4. Results

The trained LSTM model's predictions were plotted against the actual stock prices. The visualization showed that the model was able to capture the general trend of the stock prices, although some deviations were observed.

5. Discussion

The results indicate that LSTM networks can effectively model the temporal dependencies in stock price data. However, the accuracy of the predictions can be affected by various factors, such as the choice of hyperparameters, the quality of the data, and the inherent volatility of the stock market.

6. Conclusion

This study demonstrates the potential of LSTM networks for stock price prediction. While the model showed promising results, further research is needed to improve its accuracy and robustness. Future work could explore the use of additional features, such as trading volume and sentiment analysis, to enhance the model's predictive capabilities.

7. References

- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669.