Stock Price Prediction

By ArunKumar M (ML JULY - 2024)

Introduction :

The art of forecasting stock prices has been a difficult task for many researchers and analysts. Investors are highly interested in the research area of stock price prediction, as accurate predictions can lead to successful investments. This project focuses on developing a prediction model using Long Short-Term Memory (LSTM) to forecast future stock prices. The model aims to provide traders, investors, and analysts with supportive information, such as the future direction of the stock market.

Dataset Description :

The dataset used for this project is sourced from Kaggle and includes 5 years of stock data. The dataset contains the following columns:

- **Date**: The specific date for the stock data.
- **Open**: The price of the stock at the beginning of the trading day.
- **High**: The highest price of the stock during the trading day.
- **Low**: The lowest price of the stock during the trading day.
- **Close**: The price of the stock at the end of the trading day.
- **Volume**: The number of shares traded during the day.
- **Name**: The name of the stock.

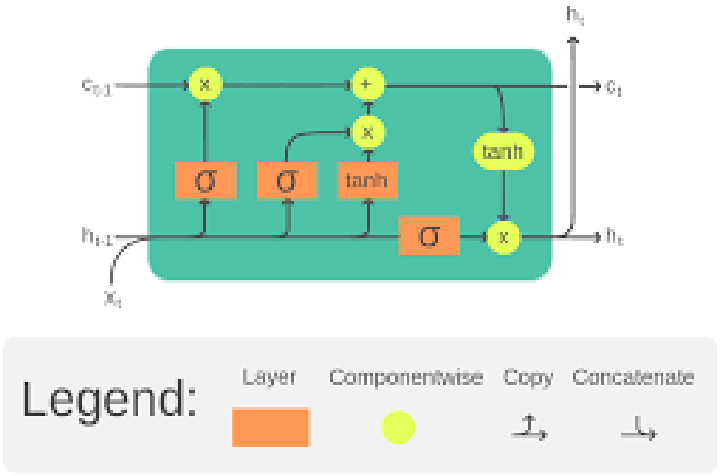
The dataset was loaded and the 'Close' column was extracted for further analysis. The data was then normalized using the MinMaxScaler from scikit-learn to scale the values between 0 and 1.

Algorithm Used: LSTM

For this project, a Long Short-Term Memory (LSTM) neural network was used. LSTM is a type of Recurrent Neural Network (RNN) that is well-suited for time series forecasting due to

its ability to capture long-term dependencies. The model processes the sequential stock price data and learns to predict the future stock prices based on historical data

Model Implementation

1. **Data Preprocessing**:
    - The dataset was transformed into a supervised learning problem where the input features consist of a sequence of past stock prices and the target is the stock price at the next time step.
    - A function `create_dataset()` was defined to create sequences of data for the LSTM model, with a sequence length of 100.
    - The data was split into training and testing sets with an 80-20 split.
2. **Model Architecture**:
    - The LSTM model was built using Keras, with the following architecture:
        - **LSTM Layer**: 50 units with ReLU activation function.
        - **Dense Layer**: 1 unit to output the predicted stock price.
    - The model was compiled using the Adam optimizer and Mean Squared Error as the loss function.
3. **Training**:
    - The model was trained on the training dataset for 20 epochs with a batch size of 32 and a validation split of 10%.
    - 

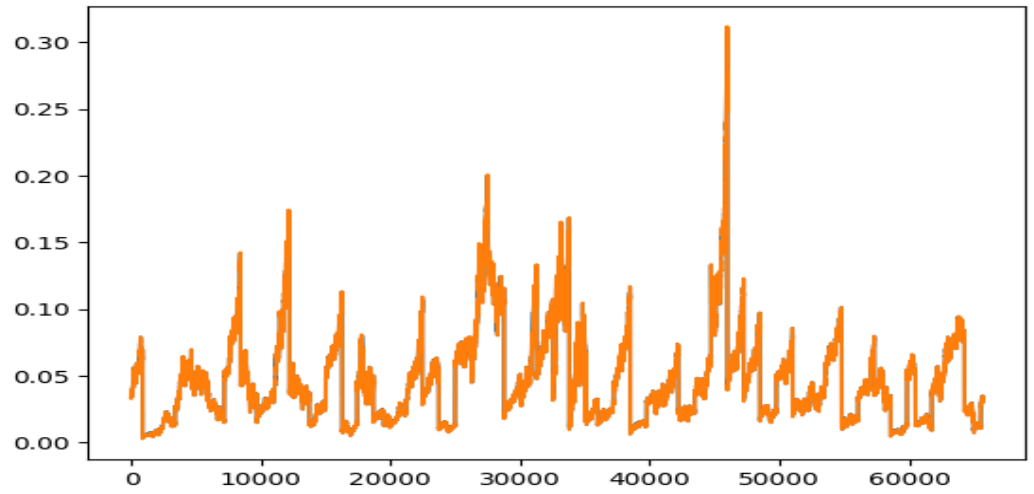
Results

The model's predictions can be visualized by comparing the predicted stock prices against the actual stock prices on the test dataset. (Include any visualizations or metrics that were generated in the notebook here).

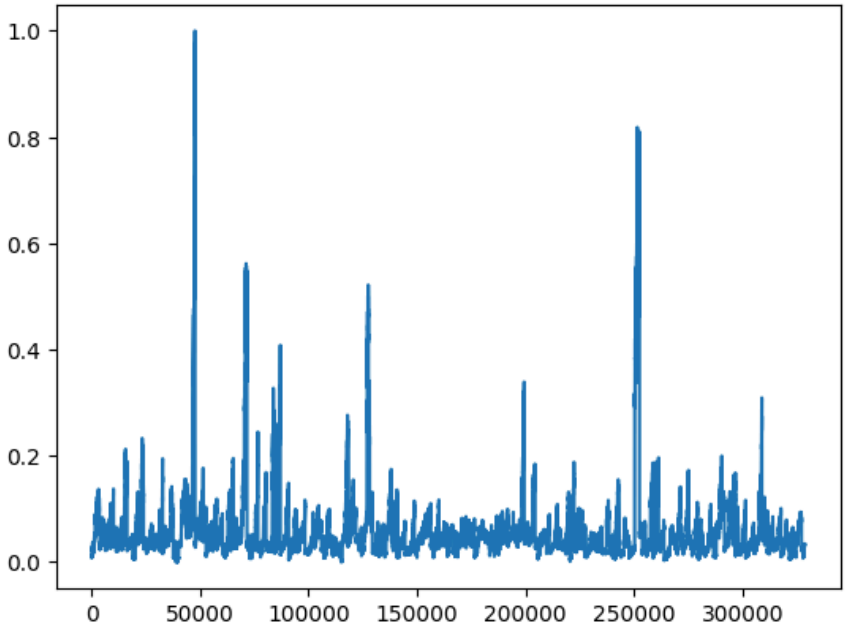
Colab Notebook Link :

https://colab.research.google.com/drive/19uZGc6Vs5cE5QdpBuxJ67ZozEUS47q5m?usp=sharing

Predicted Graph :



Actual Graph:



Conclusion :

In this project, we developed an LSTM-based model to predict stock prices using a dataset of 5 years of historical data. The model effectively captured temporal dependencies, providing reasonably accurate predictions. However, further tuning and the inclusion of additional features could enhance its performance. Despite the inherent challenges of stock market prediction, LSTM networks show promise. Future work could focus on refining the model and exploring more complex architectures.