Stock Price Prediction Using Deep Learning Models

Abstract

This project investigates the prediction of stock prices using various deep learning models: LSTM, CNN-LSTM, Bidirectional LSTM, and Transformer models. Using the stock price data of Danaos Corporation (DAC), these models are trained and evaluated to determine their performance in forecasting future stock prices. The evaluation metrics include Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), with results visualized through comparative plots to highlight the strengths and weaknesses of each model.

Introduction

Stock price prediction is a critical task in the financial industry, leveraging historical data to forecast future price movements. Accurate predictions can inform investment decisions and risk management strategies. Deep learning models have gained popularity due to their ability to capture complex patterns in sequential data. This study compares three advanced deep learning architectures to determine their effectiveness in predicting stock prices.

Data Collection and Preprocessing

Data Collection

The stock data for Danaos Corporation (DAC) was collected from Yahoo Finance, covering the period from February 21, 2014, to February 21, 2024.

Data Preprocessing

1. Normalization: The closing prices were scaled to a range between 0 and 1 using MinMaxScaler to normalize the data, which is crucial for the performance of neural networks.
2. Splitting Data: The dataset was divided into training and testing sets, with 80% of the data used for training and the remaining 20% for testing.
3. Sequence Creation: To train the models, the data was formatted into sequences. For a given day t, the model was trained to predict the stock price at $t+1$ based on the previous n days' prices.

Models

LSTM (Long Short-Term Memory)

LSTM networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problems. This characteristic makes LSTM well-suited for time-series forecasting tasks such as stock price prediction.

CNN-LSTM

The CNN-LSTM model leverages convolutional layers to extract spatial features from the input sequences, followed by LSTM layers to capture temporal dependencies. This hybrid approach aims to combine the strengths of both CNNs and LSTMs.

Bidirectional LSTM

The Bidirectional LSTM model reads input sequences in both forward and backward directions, providing a more comprehensive understanding of the data by considering both past and future context within the sequence.

Transformer Model

The Transformer model uses self-attention mechanisms to weigh the influence of different parts of the input sequence, making it effective for capturing long-range dependencies. Its parallel processing capability also allows for faster training compared to traditional RNNs.

Results

Model Evaluation

The models were evaluated based on Mean Squared Error (MSE) and Root Mean Squared Error (RMSE)

Comparative Analysis

The comparative plots of predicted versus actual stock prices for each model demonstrate that while all models show promising results, the Transformer model achieves the lowest MSE and RMSE, indicating its superior performance in capturing long-term dependencies in stock price data.

Conclusion

This project compared the performance of LSTM, CNN-LSTM, Bidirectional LSTM, and Transformer models for stock price prediction. The results demonstrate that while all models are capable of capturing the complex patterns in stock price data, the Transformer model outperforms the others, achieving the lowest MSE and RMSE. This suggests that the self-attention mechanism of the Transformer is particularly effective for this type of time-series prediction.

Future research could explore the integration of additional features such as trading volume, market indices, and macroeconomic indicators to further improve prediction accuracy. Additionally, the development of hybrid models that combine the strengths of multiple architectures could be an interesting avenue for further investigation.