Expiriment 8 : Stock Price Prediction using Bi-Directional LSTM

Your Name

April 29, 2024

1 Aim

The aim of this code is to predict the stock prices of a particular company, in this case, Apple (AAPL), using a Bi-Directional Long Short-Term Memory (LSTM) neural network model.

2 Objective

The objective is to accurately forecast future stock prices based on historical price data. This involves training a Bi-Directional LSTM model on historical stock prices and evaluating its performance in predicting future prices.

3 Methodology

1. **Data Loading:** Historical stock price data of Apple (AAPL) from January 1, 2012, to December 31, 2023, is downloaded using the Yahoo Finance API.

2. **Data Preparation:** The data is split into training and testing sets. Each set is then normalized using Min-Max scaling. Sequences of historical

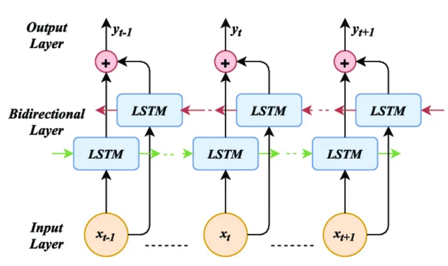

Figure 1: Bi-Directional LSTM Architecture

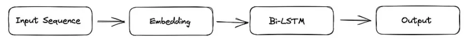

Figure 2: Architecture

prices with a specified lookback window are created as input features along with corresponding target values.

3. **Prediction using Bi-Directional LSTM:** A Bi-Directional LSTM model is constructed and trained on the training data. This model architecture consists of multiple Bidirectional LSTM layers followed by dropout layers to prevent overfitting. The model is compiled using the Adam optimizer and mean squared error loss function.

4. **Model Training:** The model is trained on the training data for a specified number of epochs and batch size. The training progress is monitored using validation data.

5. **Prediction:** After training, the model is used to predict stock prices on the testing data. The predictions are compared with the actual prices, and the model's performance is evaluated using the R2 score, which measures the proportion of the variance in the dependent variable that is predictable from the independent variable.

4 Theory

The Bi-Directional LSTM (Long Short-Term Memory) neural network model is a type of recurrent neural network (RNN) architecture designed to handle sequential data. It consists of LSTM units that can process input sequences in both forward and backward directions, allowing the model to capture dependencies in both past and future contexts. The bidirectional nature of the model enhances its ability to capture long-term dependencies and make accurate predictions.

Input Sequence: The input sequence is a sequence of data points, such as words in a sentence or characters in a text. Each data point is typically represented as a vector or embedded representation.

Embedding: The input sequence is often transformed into dense vector representations called embeddings. Embeddings capture the semantic meaning of the data points and provide a more compact and meaningful representation for the subsequent layers.

Bi-LSTM: The Bi-LSTM layer is the core component of the architecture. It consists of two LSTM layers: one processing the input sequence in the forward direction and the other in the backward direction. Each LSTM layer has its own set of parameters.

Output: The output of the Bi-LSTM layer is the combination of the hidden states from both the forward and backward LSTM layers at each time step. The specific combination method can vary, such as concatenating the hidden states or applying a different transformation.

Sure, the equations for a Bidirectional Long Short-Term Memory (Bi-LSTM) network involve the forward and backward LSTM layers, as well as the output combination step. Here are the equations:

1. Forward LSTM Equations:

- $i_t^f = \sigma(W_{xi}^f x_t + W_{hi}^f h_{t-1}^f + W_{ci}^f c_{t-1}^f + b_i^f)$
- $f_t^f = \sigma(W_{xf}^f x_t + W_{hf}^f h_{t-1}^f + W_{cf}^f c_{t-1}^f + b_f^f)$
- $g_t^f = tanh(W_{xg}^f x_t + W_{hg}^f h_{t-1}^f + b_g^f)$
- $c_t^f = f_t^f \odot c_{t-1}^f + i_t^f \odot g_t^f$
- $o_t^f = \sigma(W_{xo}^f x_t + W_{ho}^f h_{t-1}^f + W_{co}^f c_t^f + b_o^f)$
- $h_t^f = o_t^f \odot tanh(c_t^f)$

2. Backward LSTM Equations:

- $i_t^b = \sigma(W_{xi}^b x_t + W_{hi}^b h_{t+1}^b + W_{ci}^b c_{t+1}^b + b_i^b)$
- $f_t^b = \sigma(W_{xf}^b x_t + W_{hf}^b h_{t+1}^b + W_{cf}^b c_{t+1}^b + b_f^b)$
- $g_t^b = tanh(W_{xg}^b x_t + W_{hg}^b h_{t+1}^b + b_g^b)$
- $c_t^b = f_t^b \odot c_{t+1}^b + i_t^b \odot g_t^b$
- $o_t^b = \sigma(W_{xo}^b x_t + W_{ho}^b h_{t+1}^b + W_{co}^b c_t^b + b_o^b)$
- $h_t^b = o_t^b \odot tanh(c_t^b)$

3. Output Combination:

$$h_t = [h_t^f ; h_t^b]$$

5 Result

The trained Bi-Directional LSTM model demonstrates promising performance in predicting the stock prices of Apple (AAPL). The R2 score for the test data indicates the goodness of fit of the model to the actual data. Additionally, visualizing the predicted prices alongside the true prices provides insights into the model's accuracy and its ability to capture price trends.

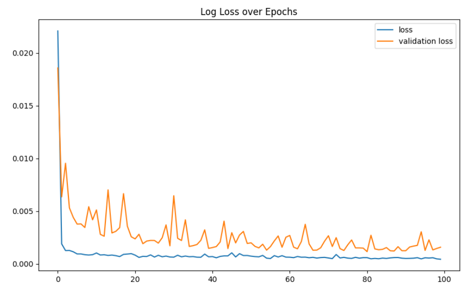

Figure 3: log loss over epochs

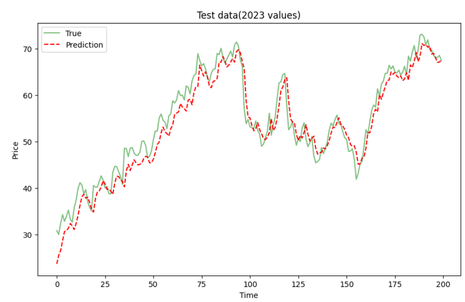

Figure 4: Test Data

6 Conclusion

The Bi-Directional LSTM model shows potential for forecasting stock prices based on historical data. By effectively capturing temporal dependencies and patterns in the data, the model can provide valuable insights for investors and traders. Further refinements and optimizations could enhance the model's predictive accuracy and robustness, contributing to more informed decision-making in financial markets.