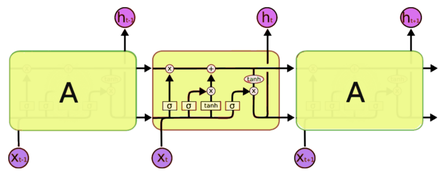
Stock Price Prediction: Deep Neural Network LSTM

Alexander Khan Applied Math and Economics Pietro Candiani Pure Math and Computer Science
Supervised by: Arafatur Rahman Financial Mathematics
Florida State University

Introduction

For a while now, many machine learning algorithms have been developed specifically for forecasting time series data. One of the most commonly used algorithms is the Long Short-Term Memory (LSTM), a type of Recurrent Neural Network (RNN). These models are particularly good at understanding long-term patterns in time series data while ignoring the irrelevant details(noise) that often trip up standard RNNs. In this project, we explore how well the LSTM model can predict the closing prices of the S&P 500, demonstrating its effectiveness in making accurate forecasts in the stock market.

Architecture of LSTM



The above figure shows the architecture of an LSTM model. What differentiates this model from other Neural Networks, are its memory cells that selectively retain or forget information over time.

This illustrates that in order to help each neuron create a more accurate forecast, they are fed information not only from the related feature but also from the neurons before and after them in the sequence.

An LSTM is a non-linear model that operates without assuming a linear relationship between predictors and response variables. It leverages non-linearity to address cases where relationships might mistakenly appear linear, incorporating an activation function at the outset to ensure diverse model responses before proceeding through its hidden layers.This Activation Function is always non-linear. The one used for the LSTM is $tanh(x)$. The model itself has the form:

$$f(X) = \beta_0 + \sum_{k=0}^{K} \beta_k g(w_{k0} + \sum_{j=1}^{p} w_{kj} X_j + \sum_{s=1}^{K} u_{ks} A_{l-1,s})$$

Where $g()$ is the activation function. It then feeds into the output from the hidden layer. It also has implemented the memory cell as it feeds the information from the previous neurons to the current neuron, resulting in $f(X)$.

Choice of Metric

The most common metric used to check for the loss of the function in a regression model is known as **MSE**

$$\text{MSE} = \sum_{i=0}^{N} (Y_i - \hat{Y}_i)^2$$

This metric checks how close the prediction is from the true value of the predictor. However, here since LSTM mimics the movement of a time series we are not only interested in them having a small MSE but also that is able to predict the movement of the stock market, since the goal of the project is to predict the price that the stock market will be.

So, instead we choose to implement our own metric to check for the accuracy of the model by seeing if it correctly predicts the movement of the market.

Feature Selection

We collected daily stock price data for open, close and volume markers over the time period starting from 2007 until the most recent data at the point of collection for 2023. We used the Alpha Vantage API, Yahoo Finance API, and Federal Reserve Economic Data to obtain this daily historical stock price data for S&P 500, Dollar Index, Civilian unemployment rate, and Consumer sentiment index.

Data	Source	Frequency	Abbreviation
Fundamentals			
Open Price	Yahoo	Daily	..
Close Price	Yahoo	Daily	..
Macroeconomics			
Civilian unemployment rate	FRED	Monthly(Int. Daily)	UNRATE
Consumer sentiment index	FRED	Monthly(Int. Daily)	UMSCENT
US Dollar Index	Yahoo	Daily	USDX
Technical indicator			
Simple Moving Average	Alpha Vantage	Daily	SMA

Table 1. Features

- The Simple Moving Average (SMA) serves as our primary technical indicator. SMA simply calculates the average stock price over a specific time period.
- UMSCENT and UNRATE play a pivotal role in stock price prediction due to their influence on consumer behavior and economic well-being.
  - UMSCENT reflects consumer confidence, driving spending and corporate earnings.
  - High UMSCENT boosts spending and stock prices, while elevated UNRATE, signaling unemployment, can hinder spending and decrease stock prices.
- The US Dollar Index (DXY) assesses USD strength against foreign currencies, impacting trade and investment.
  - Increasing DXY indicates a robust economy, influencing market sentiment and investment choices.
- These indicators are vital for comprehending the macroeconomic landscape and forecasting S&P 500 trends.

Hyperparameters

Hyperparameters are configurations that we set to optimize the model's architecture to improve the learning process of the algorithm, thereby achieving better predictions. We tuned certain hyperparameters such as the amount of neurons, batch size, and the learning rate to optimize performance. The methodology to come to our current hyperparameter configuration was brute force trial and error. We also implemented early stopping during training to optimize efficiency and performance. This mechanism halts the learning iterations when performance begins to decline, preventing overfitting and improving overall training efficiency.

Number of Neurons	126
Batch size	10
Learning rate	0.0001
Epochs	40
Optimizer	Adam

Table 2. Model Hyperparameters

Spline Interpolation For Daily Data

$$S(x) = \begin{cases} S_1(x) = a_1 + b_1(x - x_1) + c_1(x - x_1)^2 + d_1(x - x_1)^3 & \text{for } x_1 \le x \le x_2 \\ S_2(x) = a_2 + b_2(x - x_2) + c_2(x - x_2)^2 + d_2(x - x_2)^3 & \text{for } x_2 \le x \le x_3 \\ \vdots \\ S_n(x) = a_n + b_n(x - x_n) + c_n(x - x_n)^2 + d_n(x - x_n)^3 & \text{for } x_n \le x \le x_{n+1} \end{cases}$$

- A numerical method for constructing new data points within the range of existing data points.
- Particularly useful for creating smooth curves and filling in missing data points in data analysis.
- Involves using piecewise polynomials known as splines that pass through existing data points, called knots.
- Splines are defined in a piecewise domain, ensuring continuity and smoothness by connecting piecewise polynomials at the knots.

- Strategic method to refine data granularity, such as converting monthly data points into daily insights.
- Beneficial for analyses relying on macroeconomic data released monthly.
- Aims to reduce overfitting in models by providing a more detailed view of trends with daily data.

LSTM Results

The findings were intriguing, illustrating that the model adeptly followed the movements. We shall remark that the model's efficacy was markedly improved through the integration of noise data derived from the financial downturn triggered by the COVID-19 pandemic and the financial crisis of 2008.

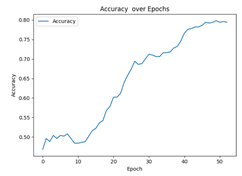


Figure 1. LSTM Price Prediction and Accuracy of Model

The LSTM's performance metric indicates that the selected features effectively predict the S&P 500's general movement, with an accuracy of about 80%. However, the model's predictive ability depends on our limited amount of chosen features and may very well be influenced by several other pertinent variables. Future studies should consider incorporating broader economic indicators or advanced technical indicators to improve predictive performance.

References

[1] Hum Nath Bhandari, Binod Rimal, Nawa Raj Pokhrel, Ramchandra Rimal, Keshab R. Dahal, and Rajendra K.C. Khatri.
Predicting stock market index using lstm.
Machine Learning with Applications, 9:100320, 2022.

[2] Jae Won Choi and Youngkeun Choi.
A study of prediction of airline stock price through oil price with long short-term memory model.
International Journal of Advanced Computer Science and Applications, 14(5), 2023.

[3] C.C. Kao, C.Y. ChiangLin, and K. C. Yang.
Applying three deep learning techniques to predicting stock price.
2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 2022.

[4] Yue Qiu, Zhewei Song, and Zhensong Chen.
Short-term stock trends prediction based on sentiment analysis and machine learning.
Soft Computing, 26(5):2209–2224, 2022.