

Stock Price Prediction Project using LSTM and RNN

Project Overview

Overview

Stock price prediction plays a crucial role in financial markets, and accurate forecasts can have significant implications for businesses, investors, and financial institutions.

Stock price prediction helps investors and fund managers make informed investment decisions. By accurately forecasting future stock prices, investors can identify potential opportunities for maximizing profits or minimizing losses. It enables them to allocate their capital wisely and adjust their investment portfolios based on predicted price movements.

Machine and Deep Learning have demonstrated their potential to enhance stock price prediction accuracy and assist investors, traders, and financial analysts in making informed decisions. By leveraging ML techniques, businesses can gain valuable insights, optimize investment strategies, and improve risk management in stock market's complex and dynamic realm.

The prediction of stock prices is a challenging task due to the inherent complexity and volatility of financial markets. Traditional methods often fail to capture the intricate patterns and dependencies present in stock price data. However, RNN and LSTM models have shown great potential in capturing temporal dependencies and making accurate predictions in various time series forecasting tasks.

In this first project of our four-part series, we will focus on implementing recurrent neural networks (RNNs) and long short-term memory (LSTM) networks specifically for the task of stock price prediction. This project serves as an essential foundation for understanding and utilizing these advanced deep learning techniques in the context of financial forecasting.

Throughout this project, we will delve into the fundamentals of RNNs and LSTMs and explore their suitability for capturing temporal dependencies in stock price data. We will learn how to preprocess and prepare the historical stock price dataset for training and testing our models.

Aim

The aim of this project is to develop a robust stock price prediction system utilizing recurrent neural networks (RNN) and long short-term memory (LSTM) networks. We will also understand the fundamentals of deep learning and its application in financial forecasting.

Data Description

The dataset used for this project consists of historical stock prices for Apple Inc. (AAPL) obtained from the Yahoo Finance API. It includes a comprehensive record of daily stock price data, capturing the opening, closing, highest, and lowest prices, as well as the trading volume for each trading day.

Tech Stack

- Language: Python
- Libraries: keras, tensorflow, statsmodels, numpy, pandas, yfinance, pandas datereader, pandas_ta

Approach

- Neural Networks Basics:
 - Review the basics of neural networks to understand their structure and functionality.
 - Building and Training Neural Networks with Keras on an Example Dataset
- Loading Time Series Data:
 - Obtain the time series data for stock prices from yahoo finance.
- Data Transformations:
 - Perform feature scaling or normalization to bring data into a consistent range.
 - Overlapping window creation for training.
- Recurrent Neural Networks:
 - Model Building and Training
 - Sequence Generation and Evaluation
- LSTMs:
 - Model Building and Training

- Sequence Generation and Evaluation
- Multivariate Input and LSTMs
 - Creating Technical Indicators
 - Creating Labels
 - Perform feature scaling or normalization to bring data into a consistent range.
 - Model Building and Training
 - Evaluation

Modular code overview:

```
├─ engine.py
├─ lib
│   ├── images
│   │   ├── decay.png
│   │   ├── lstm.png
│   │   ├── mlp.png
│   │   ├── rnn.png
│   │   ├── rnn_animation.gif
│   │   └── types_rnn.png
│   └── lstm_p1.ipynb
├─ ml_pipeline
│   ├── train.py
│   └── utils.py
├─ output
│   ├── model_lstm.h5
│   ├── model_mv_lstm.h5
│   └── model_rnn.h5
├─ readme.md
└─ requirements.txt
```

Once you unzip the modular_code.zip file, you can find the following folders.

1. lib
2. ml_pipeline
3. engine.py

4. output
 5. requirements.txt
 6. readme.md
-
1. The lib folder is a reference folder and contains the original ipython notebook as in the lectures.
 2. The ml_pipeline folder contains all the functions put into different python files, which are appropriately named. The engine.py script then calls these python functions to run the steps in one go to train the model and print the results.
 3. The output folder contains the saved models from engine.py.
 4. The requirements.txt file has all the required libraries with respective versions. Kindly install the file using the command **pip install -r requirements.txt**
 5. **All the instructions for running the code are present in readme.md file**

Project Takeaways

1. Review the basics of neural networks to understand their structure and functionality.
2. Learn how to load and preprocess time series data for stock prices, ensuring data quality and consistency.
3. Frame stock price prediction as a time series forecasting problem, considering the temporal order of the data and splitting it into training and testing sets.
4. Learn the principles and architecture of recurrent neural networks (RNNs).
5. Understand how RNNs handle sequential data and capture temporal dependencies.
6. Evaluation Metrics for Model Performance.
7. Dive into long short-term memory (LSTM) networks, a specialized type of RNN.
8. Understand the advantages of LSTMs in capturing long-term dependencies and mitigating the vanishing gradient problem.
9. Learn how to incorporate additional features or factors that may influence stock prices such as Relative Strength Index (RSI) and Exponential Moving Average

(EMA) and formulate a multivariate input problem, to enhance prediction accuracy.

10. Understand the inherent challenges and limitations of stock price prediction, including market volatility, unforeseen events, and the presence of noise in financial markets.