

Stock prediction using machine learning techniques 2023 Summer Erdos Project Executive Summary

Team: Ruojun Wang

Github: <https://github.com/Versemi/ErdosSpring2023project>

Overview:

The goal of this project is to assess and compare various prediction models for stock price forecasting. Initially, we will examine regression and time-series forecasting models to review their performance and accuracy. Subsequently, we will delve into the LSTM model, widely utilized in the quantitative finance domain, to evaluate its effectiveness and determine how it fares in comparison.

Stakeholders: Investors and Traders, Financial Analysts, Financial Institutions

KPIs: General backtesting, Volatility>Returns, Accuracy of predictions, Mean Squared Error (MSE), Return on Investment (ROI), Computational efficiency

Dataset: Yahoo Finance

Approaches: For this project, we employ three different models to assess their performance in predicting stock prices.

- Single linear regression: Analyzes the relationship between stock prices and provides insights for prediction.
- Time-series forecasting: Utilizes chronological data to identify patterns and make predictions based on historical trends.
- LSTM model: Specifically designed for sequential data analysis, LSTM considers dependencies and patterns in historical stock prices for accurate predictions.

Results:

Time-series forecasting is specifically designed to predict datasets that are structured with respect to time or dates. It leverages the inherent temporal patterns and dependencies in the data to make predictions.

When it comes to the LSTM model, it exhibits higher volatility compared to other approaches. This implies that the LSTM model has the ability to capture and detect more noise and fluctuations in the data. However, it is also more susceptible to being influenced by these fluctuations.

In terms of the accuracy of prediction, I believe that both the time-series forecasting and LSTM methods perform comparably well. They are both effective in capturing the underlying patterns and trends in the data, resulting in accurate predictions.

Pitfalls:

With additional time, I plan to consider and incorporate a more comprehensive set of evaluation metrics. Additionally, I intend to conduct further analysis to extract meaningful insights and patterns from the data. These endeavors will contribute to a more comprehensive and robust project.

