Trading At The Close

Soham Ambre (sa188), Diana Dutava (dd65) and Fernando Raverta (fr15)

Stock price prediction has been a contentious research topic, with disputes about whether stock prices are random walks and, as a consequence, unpredictable or if there is a pattern, such as mean reversion, that can be predicted [1]. The weak stock market hypothesis, which is taken as an assumption by some asset pricing methods, suggests that any information about the asset prices is already included in the prices. Consequently, this hypothesis implies that no method can achieve superior returns on a consistent basis based on historical price, volume, and returns. However, whether prices can be predicted or not may not have a global answer and may depend on the efficiency of each market and the asset itself [2].

This project, based on the “Trading At the Close” Challenge on Kaggle [3], explores stock price prediction in a focused scenario, specifically within the Nasdaq Market. The Nasdaq has a unique procedure for determining the closing price of stocks that includes releasing bid information 10 minutes before the trading day concludes [4]. The aim of this project is to assess if machine learning can help in predicting the closing price for the assets based on the information released.

The project is structured as follows:

1. **Exploratory Analysis:** We aim to conduct a thorough exploratory analysis to grasp the fundamental characteristics of the dataset. Key inquiries include investigating the magnitude of price changes in the last 10 and 5 minutes and identifying potential volatility clusters.
2. **Feature Extraction:** Try different ways of feature extraction/representation for the data to enhance machine learning algorithms' performance.
3. **Classification modeling:** We plan to develop classifiers to address binary questions, such as whether the closing price will surpass the last traded price or if it will exceed a certain threshold (e.g., xx% higher).
4. **Linear Regression Modeling:** Utilizing linear regression methods, we intend to predict closing prices based on relevant features extracted from the dataset.
5. **Neural Network Approaches:** We aim to explore the effectiveness of neural networks of varying complexities in predicting closing prices.

We anticipate encountering several challenges, including the computational demands of training models with a sufficient number of parameters to achieve accuracy in stock price prediction, given that our dataset exceeds 1 million rows. Additionally, identifying effective methods for representing data to improve predictive performance presents another significant hurdle.

By the project's conclusion, our objective is to gain a solid grasp of the effectiveness of machine learning techniques in forecasting the closing prices of stocks, assess their accuracy, and formulate well-founded strategies for enhancing future predictions.

References

[1] Dilek Durusu-Ciftci, M. Serdar Ispir, Dundar Kok. Do stock markets follow a random walk? New evidence for an old question. International Review of Economics & Finance, Volume 64, 2019, Pages 165-175.

[2] Kausik Chaudhuri, Yangru Wu. Random walk versus breaking trend in stock prices: Evidence from emerging markets Journal of Banking & Finance, Volume 27, Issue 4, 2003, Pages 75-592.

[3] Optiver. "Trading at The Close Competition".

[4] The Nasdaq Exchange. Nasdaq Closing Cross FAQ .