

Finance Club
Open Project Summer 2024

Name: Sarwade Rasika Milind

Enrollment no.: 22123039

Branch: Engineering Physics (3rd year)

Stock Sentiment Analysis using Machine Learning

Aim: The project aims to develop a sentiment analysis model to predict the movement of stock prices based on textual data from news articles, social media posts, and other sources of financial news and opinions. By analysing the sentiment expressed in these texts, the model will seek to uncover insights into investor sentiment and market sentiment, which can be valuable indicators for making informed trading decisions.

In this project I have performed the sentiment analysis using Machine Learning model (Linear discriminant Analysis) on the Infosys stock.

STEPS INVOLVED:

- 1. Collecting the data:** Collected data (headlines) using web scraping from Business Insider for Infosys for over 10 years.
- 2. Collecting data from Yahoo finance:** Collected data like Open, Close, High, Low etc.

3. **Combining the data:** Combined the stock data with headlines.
4. **Adding Features:** Subjectivity, Polarity
5. **Adding sentiment scores:** Added sentiment scores using Sentiment intensity analyzer.
6. **Splitting the data:** Splitting the data into train and test (80 and 20 percent)
7. **Applying the model:** Used Linear Discriminant Analysis.
8. **Creating a Trading strategy**
9. **Performed model evaluation :** Using appropriate metrics, such as accuracy, precision, recall, F1-score, and receiver operating characteristic (ROC) curve analysis, to assess the performance of the sentiment analysis model.

Reported Values:

Sharpe Ratio: -0.14926444808677122

Maximum Drawdown: -0.46891729600719656

Number of Trades: 165

Win Ratio: 0.28484848484848485