

Real-Time Stock Price Prediction

Amit Pingale, Shubham Patel, Vikas Kumar

Abstract

Stock market prediction is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange. The price of a stock is dependent on numerous inter-linked factors that contribute to volatility in the stock. The nexus of information corresponding to a particular asset is analyzed to determine patterns exhibited in connected services for developing trading strategies. The objective of this research paper is to investigate various branches in the corporate and government sector along with the market sentiments to predict the stock price with accuracy and minimum confidence interval.

Background

The stock market is driven mainly by uncertainty and future expectations. The price of a stock is highly complicated and is determined by a great number of variables. Global and domestic news and social media posts can hit all markets and create huge movements, all based on reaction. That makes the market difficult to predict. The analysis of this sector is complicated as there is an esoteric language used by these financial experts to explain their ideas. There are many non-statistical parameters that influence the price of stocks in the short run and this project would use correlated stocks influence on one another, sentiment analysis of influential tweets, posts and news articles as the predictor in the change to stock prices. Apart from sentiment analysis, Overseas Market/Economic Action, Economic Data, Futures Data, Buying At The Open, Midday Trading Lull, Analyst Upgrades/Downgrades, Web-Related Articles, Friday Trading should be taken into consideration for increasing the accuracy of the predictive model.

Data sources

The project requires multiple data sources:

- Text based online posts and news article for Sentiment Analysis.
- Financial data from websites that will be used to train & test machine learning algorithms.

Real-Time Data Sources

- Real time & historical stock prices
 - <https://intrinio.com/>
 - <https://polygon.io/pricing>
- APIs for news articles
 - <http://www.programmableweb.com/category/News%20Services/apis?category=20250>
 - https://developer.nytimes.com/semantic_api.json
- Twitter public REST API for public tweets from twitter
 - <https://dev.twitter.com/rest/public>
 - <https://stream.twitter.com/1.1/statuses/filter.json>

Algorithms

The following algorithms can be implemented for stock prediction:

- Long Short-Term Memory
- Recurrent Neural Network
- Gradient Boosting Machine
- Random Forest Classification
- Markov Chain
- Naive Bayes Classifier
- Kalman Filter

Model Construction

• Classification Model for Twitter Sentiment Analysis

A dictionary of keywords specific to the target stock is created and merged with existing keywords used for sentiment analysis of the tweets. The revised dataset is trained using Naive Bayes Classification algorithm for finding the polarities of tweets categorized into positive, negative or neutral.

- **Regression Model for Stock Prediction**

Transfer learning is used to speed up training and improve the performance of deep learning model. Initially, the model is trained on historical data using the LSTM algorithm. The pre-trained model can then be used as the starting point for the model to be trained on real-time data gathered using spark streaming framework. The model may need to be adapted or refined on the input-output pair data available.

Model Evaluation Metrics

- **Classification Model**

- Confusion Matrix
- Precision
- Recall
- F1
- ROC Curve
- Gini Effect
- AUC - ROC
- Gain & Lift Chart

- **Regression Model**

- Variance
- R2
- MSE
- RMSE

Resources & References

1. <https://pythonprogramming.net/finance-programming-python-zipline-quantopian-intro/>
2. <https://medium.com/mlreview/a-simple-deep-learning-model-for-stock-price-prediction-using-tensorflow-30505541d877>
3. <https://arxiv.org/ftp/arxiv/papers/1603/1603.00751.pdf>
4. Sushree Dasa, , Ranjan Kumar Beheraa, , Mukesh kumarb , Santanu Kumar Ratha,” Real-Time Sentiment Analysis of Twitter Streaming data for Stock Prediction”.

5. Guanting Chen [guanting] ¹ , Yatong Chen [yatong] ² , and Takahiro Fushimi [tfushimi] ³ "Application of Deep Learning to Algorithmic Trading".
6. <https://acadgild.com/blog/streaming-twitter-data-using-kafka>
7. <https://www.sciencedirect.com/science/article/pii/S1877050918308433>