Sentiment Analysis

Sponsor: Instat Research
Mentor: Anthony Abry
By Hunter Grubenhoff, Wesley Thorsen, Wyatt Fraley, and Matthew Green

Background

InStat Research helps model, develop and implement trading strategies for the financial markets. They use Amibroker technical analysis software to make these predictions. Due to the massive increase in social media data being generated, our mentor wanted to know if a predictive model for stock prices could be built from that data. We used various sentiment analysis techniques to determine the public sentiment related to companies that are publicly traded and provide that information to users to assist in trading strategies.

Goals

- Develop a sentiment analysis algorithm with a high accuracy rating.
- Allow users to make queries based on whatever search terms they want.
- Allow users to save queries for sentiment data.
- Provide an easy to use interface.
- Have a remote service that handles the queries, calculates and interprets the sentiment data, and passes it on to Amibroker.

Technologies Used

- Naive Bayes: Bag of words sentiment classifier.
- Fast Text: Linear classifier with n-gram context windows.
- Amibroker: Software used to simulate stock trading strategies and other analysis.

User Interface

- First Prototype vs. Final Prototype
- Allows users to query sentiment, upload new config files, or view config files based on username.

Future Work

- Implement password protection so users can keep their sentiment strategies secret.
- Analyze how Twitter’s increase to 280 characters would affect the accuracy of sentiment prediction.
- Sentiment algorithm could be improved a few percentage points with more fine tuning.
- Increase the variety of data sources used to capture sentiment.
- Increase the ease of interaction between app and databases.

Architecture

- Real time Tweets are fed through a stream processor to a database
- On a request to the RESTful Web API, tweets are retrieved for analysis
- Naive Bayes + Fast Text to predict sentiment
- Amibroker is used to analyze results
- The user then uses this to help predict changes in stock value.

Glossary

- Naive Bayes: Bag of words sentiment classifier.
- Fast Text: Linear classifier with n-gram context windows.
- Amibroker: Software used to simulate stock trading strategies and other analysis.

Acknowledgements

Special thanks to our mentor Anthony Abry for guiding us through the project, and Aimee Raley for giving us the banner, and professor Behrooz Shirazi for getting us to the end.

Team QWERTY