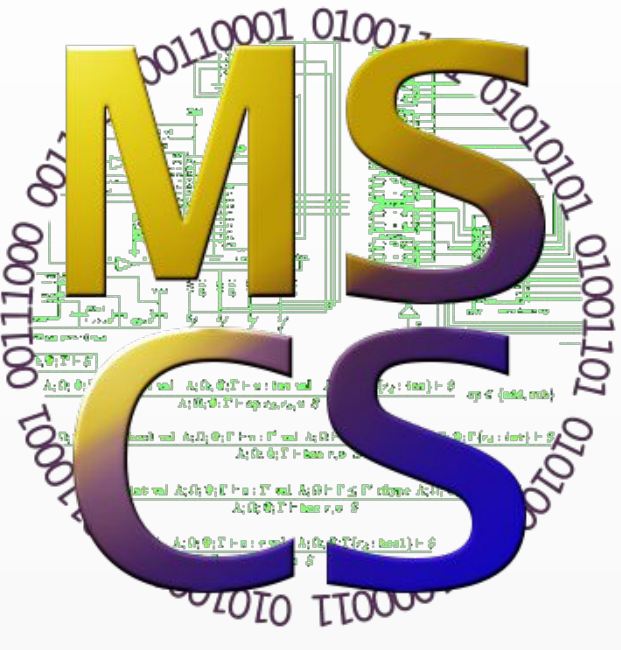




Stock Market Forecasting Through Social Media and News Analysis

Dawson d'Almeida



Background

The stock market has a major influence on the national economy. Prediction models can decrease volatility of investments and dangers of economic recessions.

Purpose

- Stock market witnesses \$77.662 trillion dollars traded worldwide yearly [1]
- Social Media is an endless supply of public data
- Emotion and sentiment are large influences on investors and stocks

Progress

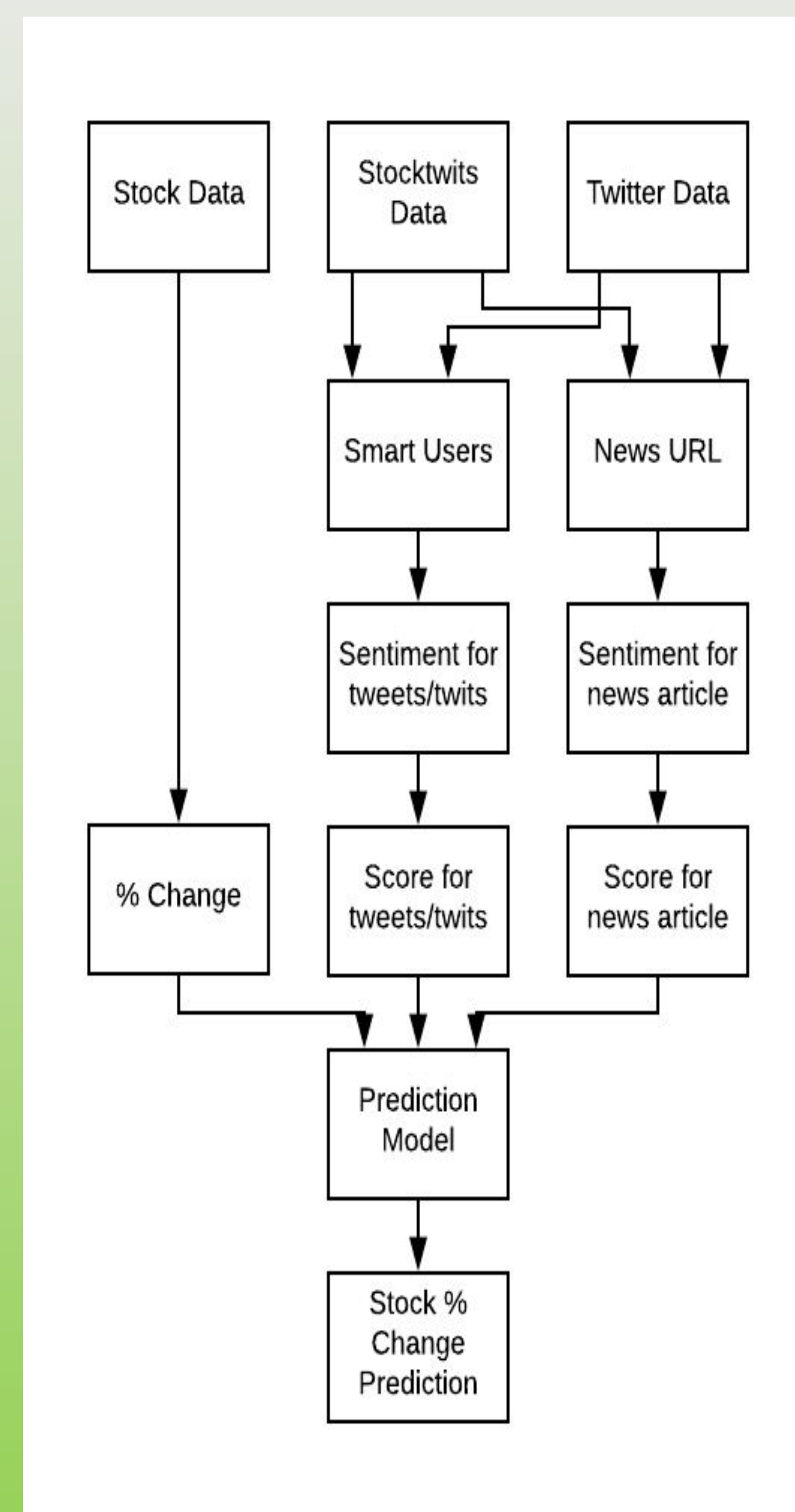
- URL content is extracted by textblob (Natural Language Processing tool)
 - Textblob provides polarity and subjectivity score from 0 to 1
- URL score is combined with twit sentiment score for prediction model

Future Work

- Use twitter data in the model with python package Tweepy
- Test different methods/models of using the URL data
 - Use more URLs from different users
- Bring in independent news articles as more data
- Use other models developed last year with the URL data
- Create a sentiment analyzer specifically catered to economic vocabulary
 - Use the Loughran and McDonald Dictionary

Paper Title	Social Media	Other Data
Modular Neural Networks	Not Used	Technical/economic indexes
Tensor Based Model	Term Vectors ("Bag of words"), TF-IDF	Firm-specific and Event-specific
Correlating S&P500 Data	Tweet Counting	Not Used
Twitter Mood Model	OpinionFinder, GPOMS	Not Used
Stock Market Analysis	Multiclass Support Vector Machine	Not Used
Investment Recommendation	Aggregate Sentiment Indexes	Not Used
Social Media Mining Tech.	Net+ sentiment dictionary	Not Used
Microblob Sentiment	Bag of Words, Word Embeddings	Not Used

Model



Symbol	Old Model Accuracy	New Accuracy	Base Case
NFLX	0.4	0.56667	0.45
AMZN	0.45	0.51667	0.45
AAPL	0.35	0.53333	0.3
TSLA	0.5	0.45	0.55
BAC	0.35	0.58333	0.35
GS	0.3	0.5	0.4
INTC	0.45	0.53333	0.5
JNJ	0.45	0.45	0.5

Reference:

[1] Stocks traded, total value (current US\$). (n.d.). Retrieved from <https://data.worldbank.org/indicator/CM.MKT.TRA.D.CD>

Acknowledgements

I would like to thank Professor Praveen Madiraju for mentoring me, Joseph Coelho for his technical help and support, and Scott Coyne for the project's foundation.

This work was sponsored in part by a National Science Foundation REU Supplement to STEM+C grant #CNS-1640217, 'PUMP-CS: Preparing Urban Milwaukee for Principles of Computer Science'.

CONTACT

Name: Dawson d'Almeida
 Email: dawsond8@gmail.com
 Phone: 360-499-1852