INTRODUCTION

It is believed that public sentiment is correlated with the behavior of the stock market. In a famous paper, Bollen et al. (2010) made the claim that Twitter mood is correlated with the Dow Jones Industrial Average (DJIA), and that it can be used to forecast the direction of DJIA changes with 87% accuracy. Besides its obvious significance in investing, this surprising result challenges several fundamental notions in the social sciences, such as the efficient market hypothesis.

In this project, I verify whether the surprising results of Bollen can be reproduced and whether they can produce a profitable investment strategy. Unfortunately, I find that measuring Twitter mood does not offer an improvement over a learning algorithm that only uses past DJIA values.

DATA

In this project, I used two main datasets:

- **Yahoo! Finance:** Dow Jones Industrial Average (DJIA) values from November 02 to November 27. The data was obtained using Yahoo! Finance and includes the hourly values of DJIA.

- **Twitter:** Twitter data containing more than 15 Million tweets from November 02 to November 27. The data is about 52 GB and is available at (http://cloud.aditya11.com). The data was collected only on trading days and during the trading hours. Twitter Streaming API was used to collect all the tweets.

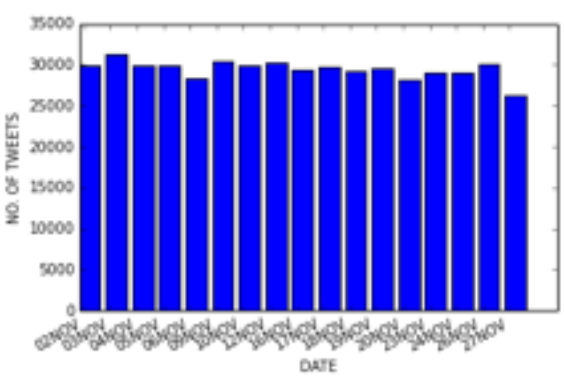

fig 1. Number of Tweets

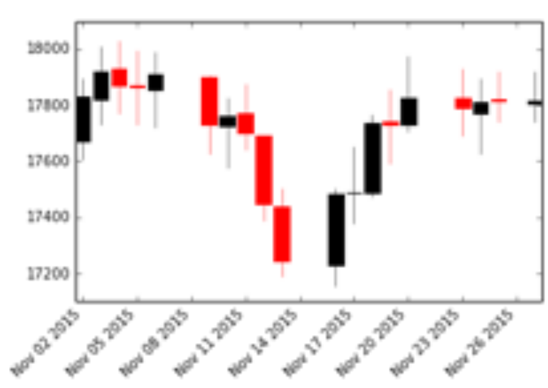

fig 2. Dow Jones Industrial Average

METHODS

The data obtained from the above sources had to be preprocessed to make it suitable for analysis.

Stocks
For the learning algorithm, I used DJIA hourly values. The data is labeled as

y = BULLISH if {opening(t + 1) > closing(t)}
y = BEARISH if {opening(t + 1) <= closing(t)}

Sentiment Analysis
We only take into account tweets that contain statements of their author's mood states, i.e. those that match the expressions "i feel", "i am feeling", "i'm feeling", "i don't feel", "I'm", "Im", "I am", and "makes me". In order to avoid spam messages and other information-oriented tweets, we also filter out tweets that match the regular expressions "http" or "www".

Sentiment analysis was an important part of my solution since the output of this module was used for learning the predictive model. AFINN word list is used to perform sentiment analysis on all the tweets.

Feature Selection
One of the main differences between my work and the original work [1] is that I try to generalize the feature set. The original author used 7 features from two sentiment analysis tools: a general positive/negative classification, and mood states in 6 dimensions of calm, alert, sure, vital, kind and happy.

Market Behavior Learning Model
Support vector machine (SVM) is used as the learning model. This model was chosen over others because SVM gave the highest accuracy compared to other models.

EXPERIMENTS

I started the experiments by performing sentiment analysis on the collected tweets. Using the AFINN word list, I found that almost 50% of the tweets were positive and 50% of the tweets were negative on any given day. So I used the number of tweets as the features.

Here is a table summarizing our results:

	precision	recall	f1-score	support
BEARISH	0.00	0.00	0.00	60
BULLISH	0.28	0.30	0.29	76
avg / total	0.15	0.17	0.16	136

The model is about 56% accurate, which is very less.

RELATED WORK

My work is based on Bollen et al’s strategy which received widespread media coverage in 2011. They also attempted to predict the behavior of the stock market by measuring the mood of people on Twitter. The authors considered the tweet data of all twitter users in 2008 and used the OpinionFinder and Google Profile of Mood States (GPOMS) algorithm to classify public sentiment into 6 categories. They cross validated the resulting mood time series by comparing its ability to detect the public’s response to the presidential elections and Thanksgiving day in 2008. They also used causality analysis to investigate the hypothesis that public mood states, as measured by OpinionFinder and GPOMS mood time series, are predictive of changes in DJIA closing values. The authors used Self Organizing Fuzzy Neural Networks to predict DJIA values using previous values. Their results show accuracy of nearly 87% in predicting the up and down changes in the closing values of Dow Jones Industrial Index (DJIA).

CONCLUSIONS AND FUTURE WORK

Given my results, the answer to the question of wether social media like Twitter can predict the stock market is currently “no”. Moreover, my algorithm achieve about a 60% accuracy by always predicting that the DJIA will go up, and therefore obviously cannot be used for constructing a portfolio that would outperform a financial instruments that follows the DJIA.

There is room for improvement and I am interested in pursuing research on this topic. The following are the potential methods of improvement that I plan to use:

Use pairs of words as features
The mutual information will increase even more with the join distribution of the words in the cluster as opposed to the distribution of the union of the words. Intuitively, this should correspond more to a contextual clustering of the words, as opposed to purely coincidental.

Use a different classification structure
The current system only considers the change as an up/down quantity and does not distinguish a 1% change in the market from a 10% change.

REFERENCES

[1] J. Bollen and H. Mao. Twitter mood as a stock market predictor.
IEEE computer, 44(10):91-94

[2] AFINN http://neuro.imm.dtu.dk/wiki/AFINN