

Turning Tweets to Trades

Sentiment Analysis for Stock Market Insight

Ellie Birbeck

supervised by Prof. Dave Cliff

Introduction

Stock markets have always reacted to news, and with the rise of social media acting as a real-time news source, lots of research has investigated whether this information can predict market movements. The use of cashtags on Twitter provides masses of specifically stock-related messages, which I will use to develop a novel approach to classification and directional price prediction.

Stock-Specific Dictionaries

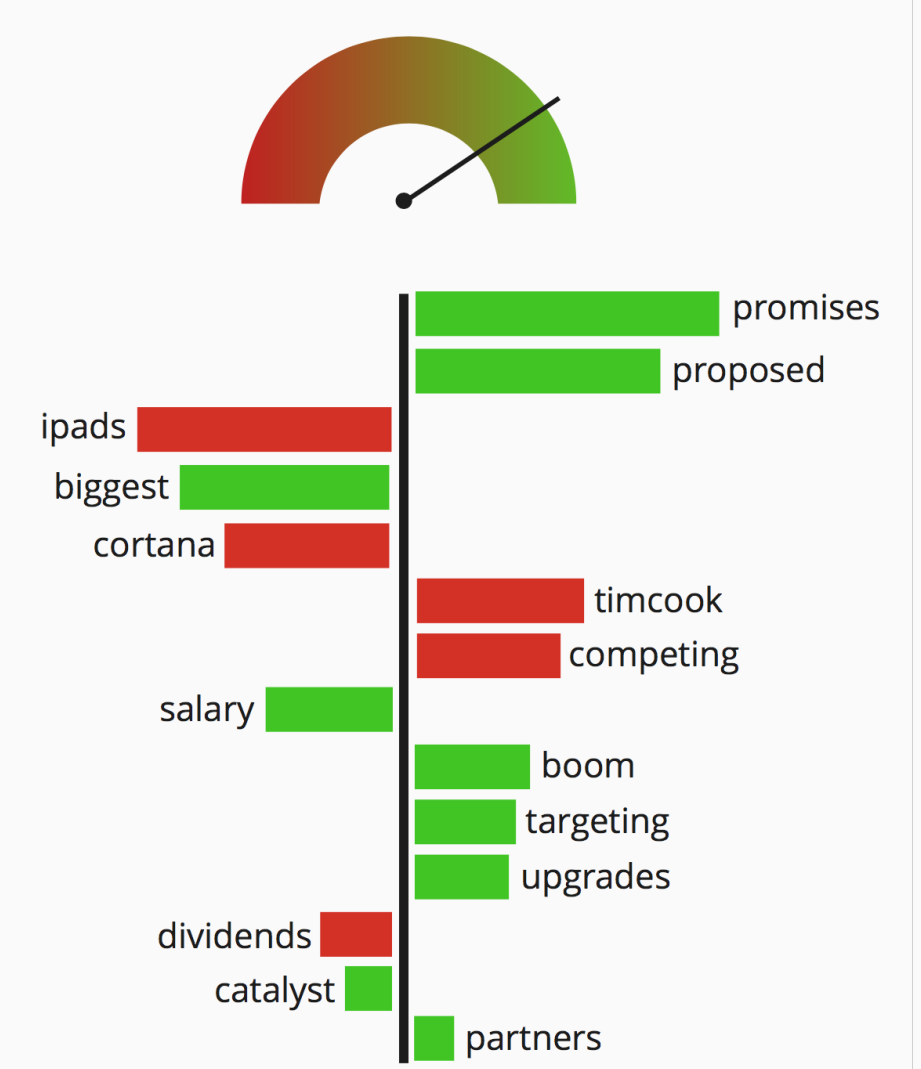
Research has shown that classifiers trained on data from one domain perform poorly when tested on another, so when considering stock-related tweets, the structure and content do not seem well suited to generic sentiment dictionaries. A financial lexicon would improve this, but I plan to test whether building dictionaries specific to a stock improves price prediction. This is achieved by using the price as the 'ground-truth' and learning from which words provide most information gain towards a change in price.

Performance Evaluation

So far I have evaluated the accuracies of several machine learning algorithms (Naive Bayes, SVM, Maximum Entropy, Logistic Regression) in predicting the direction of \$AAPL over the next hour. I will continue experiments of parameterisation and feature selection techniques (chi-squared, ROC, RFE etc.), and include further stock-related metrics as features. I will also evaluate other stocks and time-frames, and compare the overall results to predictions using a standard sentiment analysis approach.

Opinion Shift

Using the generated dictionaries, a closer analysis of why the sentiment has changed can be observed. A calculation of opinion shift will compare one corpus of tweets (e.g. today) against a previous reference (e.g. past week) and show why it is more or less bullish/bearish. Each word's change in frequency of use, and its classification of buy or sell contributes to the overall score. This will also be useful in realising when the direction predicted should be questioned, rather than using the algorithm as a black-box predictor.



Business Plan

I will develop an information subscription service targeted at casual day traders, in the form of a web app which offers them insight on the stocks they are interested in.

In addition to a dashboard of information users will be offered daily reports and real-time alerts of changes in the market, at a premium rate.

A large market exists for this type of service, and as such there are several existing competitors. The selling point of this product will be its tailored approach, in terms of the algorithm and the user experience.



University of
BRISTOL