

# Autonomous Trading Agent Implementation Using Natural Language Processing Of News Headlines

William Lyon

November 12, 2012

## Abstract

Natural Language Processing techniques are used to examine financial news headlines and generate predictions of stock price movements. News headlines from the Wall Street Journal from January 1, 2009 to November 1, 2012 are collected and paired with daily S&P 500 index returns. This information is used to train a Naive Bayes classifier and generate BUY/SELL trading signals. This trading strategy is then backtested using Wall Street Journal news headlines as a predictor for movement in the price of the S&P 500 index.

## 1 Introduction

In finance, the Efficient Market Hypothesis states that all publicly available information is reflected in financial market prices. As new information becomes available, prices adjust to take this new information into account.

### 1.1 Motivation

This tool could be used as a component of an autonomous trading agent that will make BUY/SELL decisions for trading financial instruments.

### 1.2 Literature Review

## 2 Data

Google Finance provides access to historical stock quotes as well as links to financial news stories pertaining to specific stocks. This data provides the basis for this paper. Here is some more

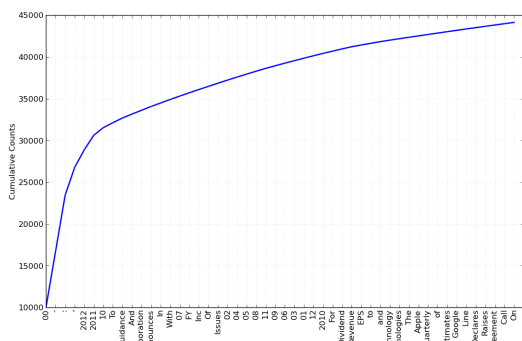


Figure 1: Cumulative frequency plot

text after the figure. The figure is somewhat floating i guess. There is even more text after the figure.

## 3 Classification Methodology

These tuples of [return, listOfDailyNews] are used as the training data for a Naive Bayes classifier.

### 3.1 Classes

### 3.2 Naive Bayes

## 4 Implementation

### 4.1 Dependencies

### 4.2 Data Collection & Munging

### 4.3 Training NLP Classifier

### 4.4 Backtesting

## 5 Evaluation

### 5.1 Accuracy

### 5.2 Most informative features

### 5.3 Precision

### 5.4 Recall

## 6 Trading Model

### 6.1 Trading Signals

### 6.2 Backtesting

## 7 Conclusions

### 7.1 Further research

## 8 References