Sentiment analysis on the stock market

Mikel Iparragirre and Aimar Ugarte

Introduction

- First exchange house established in 1611 in Amsterdam
- Open Outcry system used through most of history
- Telephone Transactions
- NASDAQ exchange opened 1971
- First Electronic conversion in 1986
- Twitter created on 2006

Stock market prediction using machine learning techniques

- Efficient-market hypothesis (EMH) says its not possible to outperform the market through stock selection or market timing.
- But technical analysis believes that stock market is strongly affected by politics, general economy, expeculations, psychology of investors, etc.
- Artificial Neural Networks (ANN) and Support Vector Regression (SVR) the most used algorithms.

Stock market prediction using machine learning techniques - Example

- There are many algorithms that use machine learning to predict the stock.
- The artificial fish swarm (AFSA) to train Radial basis function neural network (RBFNN).
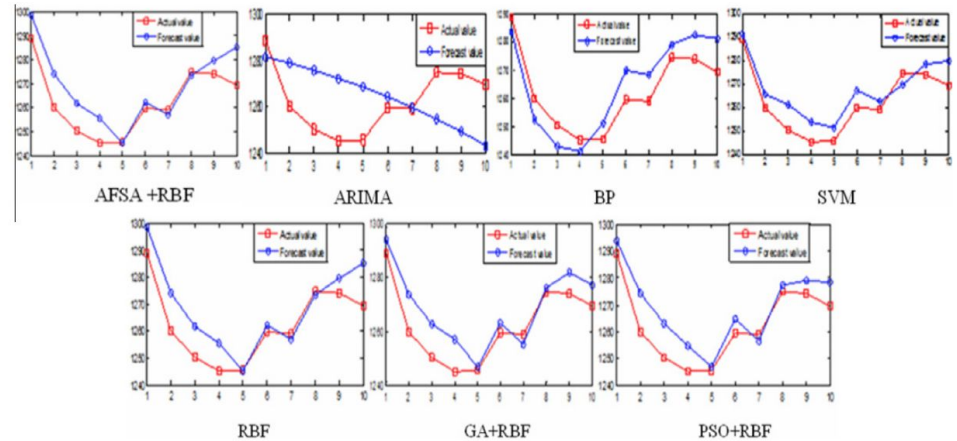- This kind algorithms uses quantitative factors but stock is affected also by non-quantitative factors too.

Fig. 5. Comparison of the forecasting results of seven discussed models.

Sentiment analysis

- Its a field of Natural Language Processing (NLP) that aims to determine the attitude of a speaker or a writer.
- Besides identifying the opinion they extract other information:
  - Polarity
  - Subject
  - Opinion holder
- There are also three categories of algorithms in the sentiment analysis field:
  - Rule based
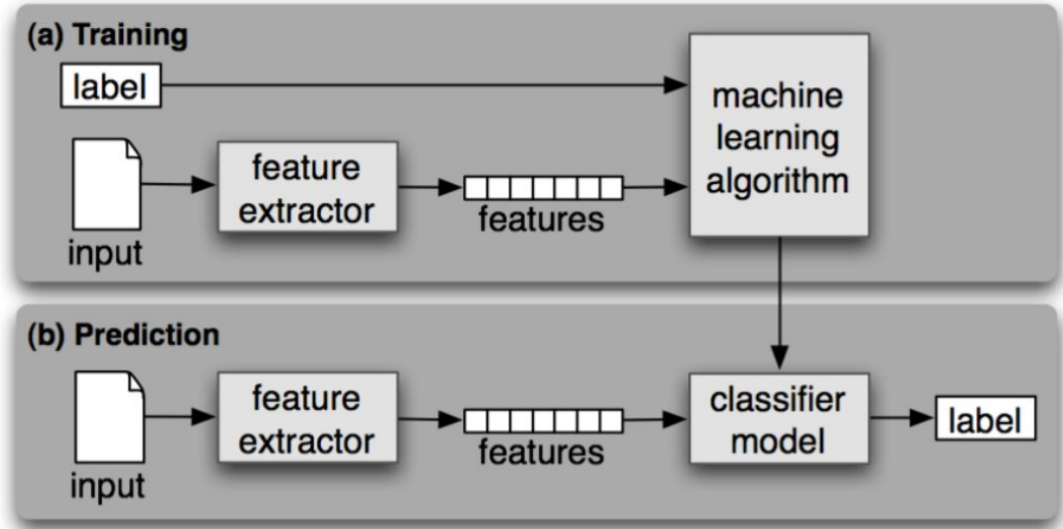 - Automatic systems
  - Hybrid

Sentiment analysis - Rule based algorithm

- Defines the rules in some scripting language.
- A basic example would be the next:
  - Define two list, negative and positive words
  - Given the text count the number of bad and good words
  - If the number of bad words is higher than its negative sentiment, otherwise it is positive sentiment
- It does not take into account how words are combined in sentences.

Sentiment analysis - Automatic systems

- Don't rely on manual crafted rules.
- The algorithm has two stages:
  - Training stage
  - Prediction stage
- On the training stage, the algorithm is fed with training input.
- On the prediction stage it categorizes the input received.

(a) Training

label → machine learning algorithm

input → feature extractor → features → machine learning algorithm

(b) Prediction

input → feature extractor → features → classifier model → label

Sentiment analysis - Hybrid

Combining the best of both worlds the methods usually improve in accuracy and precision.

Sentiment analysis on the stock market

- One of the field tha sentiment analysis is applied is the stock market.
- With the right input such as News, Tweets, Blogs
- It can be predicted if its a good buying or selling opportunity.
- Sentiment analysis made by Sentdex on the Microsoft stocks.

Stock: MSFT
OHLC (TOP) and Sentdex Sentiment Index (BOTTOM)

Twitter and the stock market, the power of twitter

- 332 million active users in 2016
- Most state leaders represented
- 20 year old fighter project in danger after a tweet
- Trump makes Toyota lose $1.2 billion dollars over a tweet
- Elon Musk loses CEO post and 20 million dollars for a tweet

Sentiment analysis on Twitter

- Influential article "Twitter mood predicts the Stock Market" (1550 citing articles) published in 2010 about prediction of DJIA using Sentiment Analysis with a reported accuracy of 86%.
- A paper inspired by it achieved accuracy of 75.56% in 2011 with 476 million tweets analysed over 6 months and correlated to DJIA.
- Through the years more Algorithms applied.
- Improvements in mood measurement and data preprocessing

Conclusion

- Clear correlation between twitter mood and stock market fluctuation established
- Capacity to predict the market confirmed
- Electronic and automatized transaction dominate the stock market
- A greater automatization is expected