

ARMY INSTITUTE OF TECHNOLOGY



Information Technology Department

PROJECT REVIEW 1

PREDICTION ON MARKET MOOD USING SENTIMENTAL ANALYSIS

Statement about the Problem:

Stock market trading is a complex process where traders aim to maximize their expected return while minimizing associated risks. With the increasing availability of digital historical records, using automated agents for stock market trading becomes of a significant interest. Different machine learning algorithms are used to circumvent the problem of defining explicit targets and tackle problems which require sequential decisions. The prediction of mood/trend of stock markets is a vital problem, where a trader wants to make a position with predefined amount of shares over a fixed time horizon in the predicted trend.

Why is the particular topic chosen?

In recent years, stock market analysis has been an important consideration for companies and industries to align their business strategies. A stock market is a place where sellers sell their company's stocks while the buyers aim to increase their worth by buying stocks. However, deciding which stock to buy and sell depends on how the company is expected to function in the future. It is very beneficial in market analysis and for the investors to choose the stock to invest in. Thus stock market analysis involves handling this volatile nature of stock data.

People use social networking sites, like Facebook, Twitter, etc. to express their opinions and views about a particular topic such as news, movie, event and remarks related to product. This information available from social network is beneficial for business analyst for mining the user opinion and predicts the mood of market and reduces the risk of investors.

Objective and scope of the project

Before investing money, it is very important for investors to predict the stock market. In today's digital world Internet based technologies such as Cloud Computing, Big Data analytics, and Sentiment analysis have changed the way we do business. Sentiment analysis or opinion mining makes use of text mining, natural language processing (NLP), in order to identify and extract the subjective content by analysing user's opinion, evaluation, sentiments, attitudes and emotions.

The forecasting of the stock market price is very important in the planning of business activities. The prediction of the stock price movement has gained a lot of interest from the researchers across various disciplines including computer science, economics, finance, statistics, etc. Analysis of the stock market is now a multidisciplinary process involving social media analysis, using insights from news headlines, crunching past data and using technologies like Machine Learning, Deep Learning, Time Series Analysis, Sentiment Analysis ,NLP, etc.

Proposed system Architecture

Here we are discussing about sentiment analysis

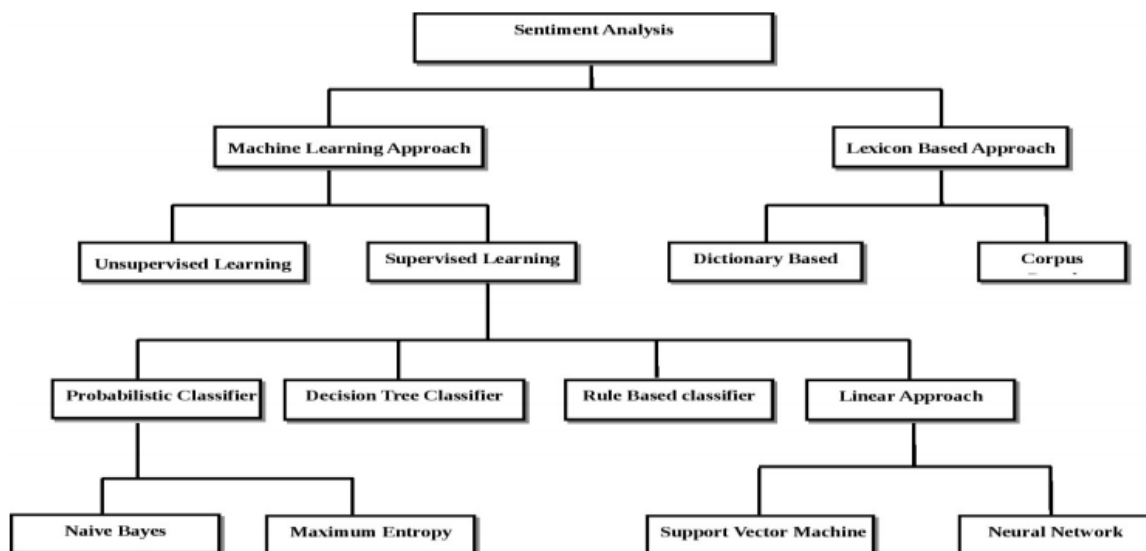


Figure1: Sentiment Classification Techniques

We are discussing here machine learning approach:

Machine Learning Approach Machine learning techniques that are applied in the field of sentiment analysis can be divided as supervised and unsupervised learning methods.

Unsupervised Learning

Unsupervised learning has no explicit target output associated with input, and it is learning through observation. The goal is to have the machine learn without giving any explicit instruction. Famous approach in unsupervised learning is clustering, in which similarities of elements in the training data is found out. Cluster similarity parameter is defined upon metrics such as Euclidean distance. K-means, Hierarchical, Gaussian mixture models, Self-organizing maps, and Hidden Markov models are some of the clustering algorithms.

Supervised Learning

Supervised learning is one that makes use of known dataset to make the prediction of output result. Supervised learning requires two sets of documents: training set and test set. For learning different properties of documents, training set is used and for evaluating the performance classifier test set is used.

1) Decision tree classifier

2) Rule based classifier

3) Probabilistic classifiers:

Probabilistic classifiers are developed by assuming generative models which are product distributions over the original attribute space or more involved spaces. While this paradigm has been shown experimentally successful on real world applications, despite vastly simplified probabilistic assumptions. Probabilistic classifier is based on the prediction of input given probability distribution. Two most important probabilistic classifiers are discussed as follows

Naïve Bayes classifier: This classifier is based on Bayes theorem of probabilistic model. In this we tried to estimate the probability of a text based on whether it belongs to positive or negative class.

Maximum Entropy classifier:-Maximum Entropy classifier is a probabilistic based classifier which belongs to the exponential model class. Principle of maximum entropy is used in this chapter and distribution having largest entropy is chosen .

Methodology

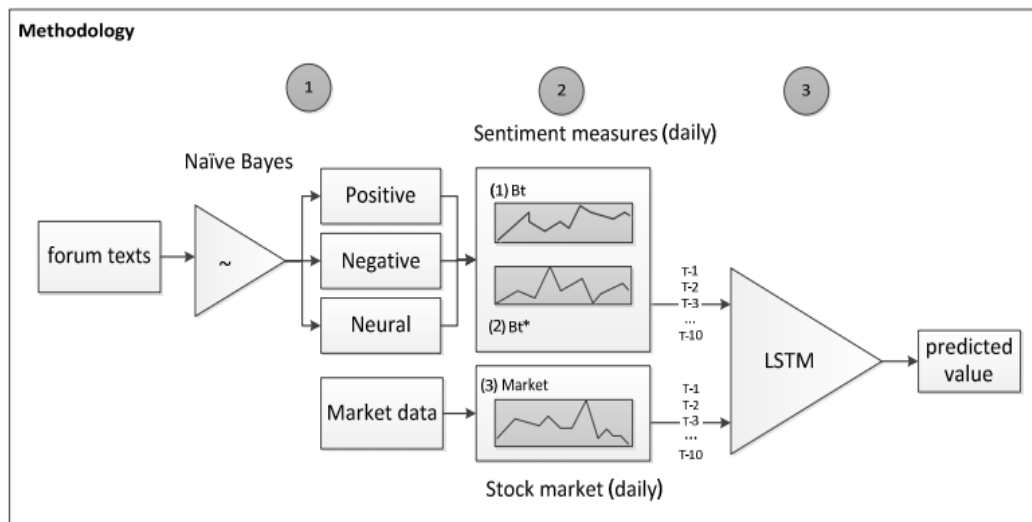


Figure 1. Methodology

We proceed in three phases. In the first phase, we employ a Naïve Bayes method to classify these posts into one of three categories: positive, negative, or neutral.

In the second phase, we construct an investor sentiment index to measure the daily mood of stock market. In the third phase, we deploy a Long Short-Term Memory model to test the hypothesis that the prediction accuracy of stock market prediction models can be improved by including measurements of investor sentiment.

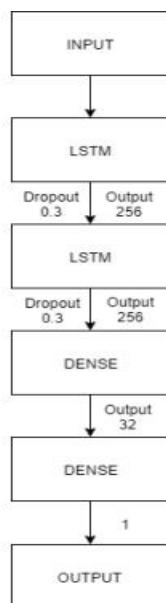


Figure 2. LSTM Layers

Hardware & Software to be used

Software:

- Python
- Jupyter Notebook

Hardware:

- Process Intel10 core processor, 2.2 GHz with Turboboost upto 3.1 GHz.
- 25 MB Cache
- Motherboard – ASRock EPC612D8A
- RAM – 128 GB DDR4 2133 MHz or more
- 2 TB Hard Disk (7200 RPM) + 512 GB SSD

What contribution would the project make ?

Use of recently introduced machine learning techniques in the prediction of stocks have yielded promising results and thereby marked the use of them in profitable exchange schemes. It has led to the conclusion that it is possible to predict stock market with more accuracy and efficiency using machine learning techniques. In the future, the stock market prediction system can be further improved by utilizing a much bigger dataset than the one being utilized currently. This would help to increase the accuracy of our prediction models.

Guide :Prof G . Walunjkar

- 1) Mayank (4330)
- 2) Payal(4336)
- 3) R Sankalp Shukla(4340)
- 4) Shikha Jha (4351)