# PAPER SUMMARY

**TITLE:** Sentiment Analysis for Effective Stock Market Prediction

**AUTHORS:** Shri Bharathi, Angelina Geetha

**PUBLISHED:** International Journal of Intelligent Engineering and Systems

**YEAR:** June 2017

**SUMMARY:** This paper attempts to design and implement a predictive system for guiding stock market investment. The approach is the combination of both sensex points and Really Simple Syndication (RSS) feeds for effective prediction. This paper claims that the sentiment analysis of RSS news feeds has an impact on stock market values. Hence RSS news feed data are collected along with the stock market investment data for a period of time. Using the algorithm for sentiment analysis, the correlation between the stock market values and sentiments in RSS news feeds are established.

**AIM OF THE PAPER:** This paper focuses on developing a successful approach to predict stock prices along with minimizing the inaccurate forecast of stock prices. Two types of results are achieved in the proposed system. One is without sentiment the moving average value is calculated and the second one is with sentiment the moving average is calculated. Both the results will show the improvements in the stock market forecasting. This paper explores the task of automatic identification of news opinions with the help of RSS news feeds and predicts the stock market movement whether goes up or down.This trained model is used for prediction of stock market rates. In this experimental study the stock market prices and RSS news feeds are collected for the company ARBK from Amman Stock Exchange (ASE). The study has shown an improvement of 14.43% accuracy prediction, when compared with the standard algorithm of ID3, C4.5 and moving average stock level indicator.

**RSS NEWS FEED:** Really Simple Syndication or Rich Site Summary (RSS) is an XML document that facilitates content syndication. It is a efficient way to read the news-related sites, Weblogs and other online publishers syndicate their content as an RSS feed to whoever wants it.With RSS feed, users can finally separate wanted information from unwanted information. Most RSS news feed contains author, title, and date information in addition to link and description. If there is positive RSS news feed then this tends to have a positive effect on stock markets and share prices rising soon after the news come out in the Open. On the other hand, negative RSS news has negative impact on stock markets and makes the stock market values go down.

**STOCK LEVEL INDICATOR:** Initially the historical prices of the selected company are

downloaded from the website. Various methods of stock level indicators are available to computing the stock value like Moving average method (it offers a smoothed line and also helps to cut down the amount of noise on price chart compared with other level of indicators). The sensex is designed to reflect the overall market sentiments.

$Sensex$ = ($sum\ of\ free\ float\ market\ capitalization$ / Base Market Capitalization ) x 100

$Free\ Float\ Market\ capitalization$ = $Share\ price(shares\ out\ standing - locked\ in\ share)$

The moving average is calculated by adding the closing price and then dividing this total by the number of time periods.

$Ft$ = ( $At-1+At-2+At-3+\cdots+ At-n$ ) / $n$
$Ft$ = Forecast for the coming period,
$At-1$ = Actual occurrence in the past period for up
to 'n' periods,
N = Number of periods to be averaged.

**NLP MODULE:** Sentiment analysis or Opinion mining is the process used to determine the attitude, opinion, emotion expressed by a person about a particular topic. Sentence level sentiment score (SSS) Algorithm is considered for finding the overall result where for every individual sentence analysis approach is applied and finally their results are summarized to provide the overall result of the document.
$Sequence\ of\ words(W) = W1 + W2 + \cdots + Wn$
S($W+ve$) = Set of Positive Sentiment words.
S($W-ve$) = Set of Negative Sentiment words.
n = Number of words.
If the score value of that sentence is positive then that sentence is considered as a positive sentence. If score value is negative then it is considered as negative sentence. If it is 0.0, then it is considered as a neutral sentence.
Dictionary based approach: This approach assigns each synset of WordNet into three sentiment numerical scores: Obj(s), Pos(s), and Neg(s) describing how Objective, Positive and Negative the terms contained in the synset are.

**COMBINING NLP MODULE WITH STOCK INDICATOR MODULE:** To predict the stock market, the results of both sentiment analysis and Sensex are combined and analysed.

| Sentiment Analysis Result | Sensex-Moving Average Result | Final-Result Prediction |
|---|---|---|
| Positive | Positive | Positive |
| Positive | Negative | Neutral |
| Negative | Positive | Neutral |
| Negative | Negative | Negative |

**RESULTS:**

Without sentiment value moving average produces the accuracy of 64.32% and with sentiment value the moving average produces 78.75% with improvement of 14.43%. In the previous algorithm ID3 and C4.5 the accuracy is less than the proposed algorithm. The accuracy of the measure is defined by precision and recall.

$Precision = TP / (TP+FP)$

$Recall = TP / (TP+FN)$

| ARBK Company | ID3 | C4.5 | Moving Average | Moving Average+ Sentiment Analysis |
|---|---|---|---|---|
| Total Instances | 499 | 499 | 499 | 499 |
| Correctly Classified | 233 | 237 | 321 | 391 |
| Precision % | 46.69% | 47.49% | 64.32% | 78.75% |