IJCSN International Journal of Computer Science and Network, Volume 4, Issue 4, August 2015
ISSN (Online) : 2277-5420 www.IJCSN.org
Impact Factor: 0.417

Sentiment Score based Algorithmic Trading

[1] **Sukesh Kumar Ranjan,** [2] **Abhishek Trivedi ,** [3] **Dharmveer Singh Rajpoot**

[1,2,3] Department of Computer Science and Engineering / Information Technology,
Jaypee Institute of Information Technology, Sector-62 Noida, Uttar Pradesh, India

Abstract - Forex trading is one of the most volatile markets which basically depend on currency exchange. There are so many situations like disasters, political activities, etc. which affects the exchange rate of currency. Prediction of Price or Exchange rate can be done either via technical approach or fundamental approach, but there are some flaw exists in both the approaches, so we propose a sentiment scored based algorithmic approach which is a hybrid model that overcomes the shortcomings of existing approaches. The experimental results show that the proposed method gives the better accuracy as compare to traditional approaches.

Keywords - *Trading, Sentiment, Profit, Loss, Risk etc.*

1. Introduction

Forex trading is buying and selling of currencies in global market. It is traded more than 3.5 trillion dollar per day and consider to be high volume market than other trading markets.Forex trading always comes in circle of attraction for all investors. It has high volatility which eventually attracts traders more than any other investment sectors. There are basically two styles of trading one is fundamental trading and other is technical trading.

1.1. Fundamental Trading

In the Fundamental trading, any trader analyses the market on the basis of news sentiment and then trader will react to the situation and take trade related decisions. This approach cannot be considered as computational model for market analysis because the accuracy is variable which depends over experience of trader. Emotions and Ego also involves in decision making which should be avoided at time of trading Maintaining the Integrity of the Specifications.

1.2. Technical Trading

In the Technical trading, trader rely on trend and indicator formed by data where trend is used to decide the behavior of any trader. It also has some drawbacks like after the occurrence of any news its effect will be observable after a long time so the length of frame used for observation usually decrease the opportunity for profitable trade. However, it has some advantages over fundamental trading such as It is free from personal emotion and ego of the trader which deeply effects the decision making, and makes the decision more rational.

1.3. Related Terms

In this part we are providing necessary backgrounds for better understanding of algorithmic trading and related terms. There are few technical terms which is necessary for deep understanding of any trading model these are following**:**

1.3.1 Algorithmic Trading

Algorithmic Trading is a kind of trading which is performed by pre- programmed instructions, which is made for the purpose of trading automatically without putting orders manually. We make strategies with the help of trend and momentum indicators. Then trade according to some specific condition for taking trade decisions. In order to make complete strategies for algorithmic trading we need three parameters namely.

(a) Take Profit Level: Take profit level is existing strategy level which triggers after earning maximum profit attained by reducing chances of loss. In figure 1, we have consider 1.16 level as Take profit level, which signifies that algorithmic strategy will automatically stopped at specific level of 1.16 in figure 1 and 1.245 in figure 2.

(b) Stop loss Level: Stop loss is again a measurement where trader can stop his strategy in case of worst scenario. It means after this

benchmark it will stop trade. Stop loss is better for trading as if we are getting loss in open orders then if it reach stop loss target then we can lose only small amount of money rather than large amount of money. It also open the way of opening new orders according to new trade decisions. In figure 1, we have consider 1.19 level as Stop Loss level, which signifies that algorithmic strategy will automatically stopped at specific level of 1.19 in figure 1 and 1.235 in figure 2.

(c) Trade Decisions: In this section, trader simply decide, whether trader have to buy or sell. Forex trading is mostly based on probabilities of success and failures. Most of the traders try to open that orders that have probabilities of making profits and easily reach their take profits target. So we mostly rely on trend or direction of moment in Forex trading. If trend is down then it is known as downtrend and if trend is up then it is known as uptrend. First of all traders remove odds by finding direction of moment, if trend direction is up then we only open buy orders and if trend direction is down then we only open sell orders. Thus we can easily make trade decision on the basis of trend direction. Figure 1 signifies trend is down and figure 2 signifies trend is up.

Figure 1 Downtrend

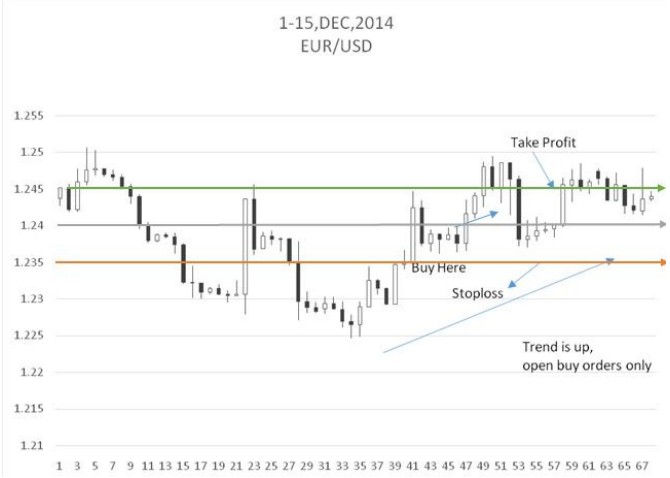

Figure 2 Uptrend

In this paper, we propose a hybrid approach which combine the concept of fundamental trading and technical trading with the help of sentiment score, which further classifies the news into six different classes. We are considering news as a trigger point for any calculation and analysis in our model. We are also overcoming the drawback of technical approach by reducing time needed for trend analysis in technical trading. After occurrence of any news we have to calculate the sentiment score value which will be helpful to identify sentiment and trend in a short time duration. Rest of the paper organized as follow in section 2 we explain related work has been done in this field . In section 3 we explain about our proposed model with all related notations and methods. In section 4 we have mentioned all experimental results related to this model and its comparison with other benchmarks. In section 5 we have indicated the conclusion and future work for this model.

2. Related Work

Lot of researches and work has been already done by various researchers in the field of market analysts. They also use financial news as a trigger point for any observation and analysis. Some details are following:

Xiaodong et. al. [1], explain how ELM (Extreme learning Machine) can be used to predict the financial markets signals. This paper also includes the comparison of SVM and ELM and founded both approaches gives higher accuracy but ELM is faster than SVM in predicting the market scenario. These are some steps (Labeling, modeling and evaluating) that are followed in order to complete the market prediction. Arne Thorvald et. al. [2], explain how financial news can be classified in order to

predict the stock market moments. It uses SVM along with WordNet-transformed bag-of-word for the required analysis. Thomson Reuter's corpus is used also for the analysis.

Link provided by Thomson Reuters,[3] explains how 'News Analytics' uses text scores and metadata for the classification of financial news. Which forecasts the market variables as the effect of the news. It first reprocess the texts of news and gives some scoreable indicators like Revalance and Prevailing Sentiment.

All these approaches are also useful for making trade decision but for complete algorithmic trading we also need take profit level for satisfying boundary constraints and then only we will be able to compare our model with other existing models and indicators like moving average and RSI indicator.

3. Proposed Model (Sentiment Score Based Algorithmic Trading)

As we have already mentioned that it is hybrid model of two existing approaches technical trading and fundamental trading. We have divided our proposed model in two Phases. In Phase 1, we developed old datasets for future price predictions .In Phase 2, we calculated the all three parameters for algorithmic trading with the help of Phase 1.

3.1 Phase 1

For price prediction using our model we have to prepare old data set of last 15 days for observation showing in figure 3.

3.1.1 Basic Terminology

In proposed model we are using some technical terms which is necessary to deep understanding of our model. These terms are follows:

(a) Time Frame: We have taken the time frame as four hours for any calculation regarding our model. We have taken this time frame randomly any one can take any other time frame as it is just for segmentation of prices in chunk as in figure 3.

(b) Base Point: We have taken base point as 1000. It is just for showing the relative psychology of trader and market. This will helps us to predict the market by combining the effect of all news rather than relying on single news. Any other trader can alter this value but round off and multiple of 100 is preferable for better calculation.

(c) Classification of News: If there is any news either scheduled or randomly occur then it will affect the market moment certainly. For the further analysis in this model we have to classify news in two major classes and every major classes there are three sub classes as given in table 1.

Table 1: Classification of news and its impact.

Positive Classes News	Negative Classes News
High Impact	High Impact
Medium Impact	Medium Impact
Low Impact	Low Impact

(d) Criteria for classification: Any news comes at a particular time frame't', we can calculate price change by using the following formulae

$$PriceChange = \{ClosePrice(t) - OpenPrice(t)\} * 10000 \quad (1)$$

Based on the value computed by equation no. 1, the price change will definitely lie in the following classes.

Table 2: Impact on market on the basis of price change

Price Change	Classes
Price_Change>50	Highimpact in positive sense
30<Price_change<=50	Medium impact in positive sense
0<Price_change<=30	Low impact in positive sense
-30<=Price_change<0	Low impact in negative sense
-50<=Price_change<-30	Medium impact in negative sense
Price_Change<-50	High impact in negative sense

We have taken boundaries value (0 -30, 30-50) after performing statistical testing over the chunk of 50 news of each classes and calculated average price change. According to the average price change we have chosen above described boundaries values.

(e) Relative Impact: Relative Impact will show the variation in the base point according to the news and their impact over the market. We have earlier mentioned the classes for the news, now we will add a new component which is score of the class (News classified Score). Initially we are taking Base Point as Relative Impact Score because initially there will be no news in the queue. Relative Impact can be calculated by the following formulae.

$$RelativeImpact = RelativeImpact + NewsClassifiedScore \quad (2)$$

Range of new classified score

Table 3: Relative impact relation with News classified score

News classified Score	Impact	Relative Impact
30	High impact in positive sense	Relative Impact =Relative Impact +30
20	Medium Impact in Positive sense	Relative Impact =Relative Impact +20
10	Low impact in Positive sense	Relative Impact =Relative Impact +10
-10	Low impact in negative sense	Relative Impact =Relative Impact -10
-20	Medium impact in negative sense	Relative Impact =Relative Impact -20
-30	High impact in negative sense	Relative Impact =Relative Impact -30

.

Up to now we have prepared old datasets for analyzing the market. Now we need three parameters for algorithmic trading.

3.2 Phase 2

3.2.1 Trade Decisions

According to our model we are not going to predict any decisions only based on trend analysis but also using sentiment related to the news as a component in decision making.

(a) Trend of Market: Suppose we have to predict the market at any particular frame of time then we have to calculate trend at time fame 't' by taking the differences between open price of fifteen days prior to the time slot 't' and the close price of current time frame 't-1' and we will denote the difference by delta.

$$\Delta = Close\,Price(t-1) - Open\,Price(t-15) \quad (3)$$

Table 4: Trend Direction

Δ	Trend
Δ>0	Uptrend
Δ<0	Downtrend

We have taken fifteen days as minimum time to wide higher computation in order to predict the trend. But anyone can take different value but the value should be greater than fifteen days. As Forex trading market is open for twenty days in a month so it will cover at least 75 percent market behavior in a month.

(b) News Sentiment Analysis:
If base point is less than relative impact score then it will be consider as positive sentiment otherwise negative sentiment.

Table 5: Relation between base point and Relative Impact score

Base point vs Relative Impact Score	Type of Sentiment
Base Point > Relative Impact Score	Positive Sentiment
Base Point < Relative Impact Score	Negative Sentiment

After concluding these two factors we have to take trade decisions.

Table 6: Trade decisions on the basis of Parameters

Parameters	Trade Decisions
Sentiment=Positive and Trend=Uptrend	Buy
Sentiment=Negative and Trend=Downtrend	Sell

Sentiment=Positive and Trend=Downtrend	Bubble, Trade should be avoided
Sentiment=Negative and Trend=Uptrend	Bubble, Trade should be avoided

We should avoid trades when it forms bubbles. Bubble is a situation where trend and news sentiment are opposite, high risk is associated with the trade decisions.

3.2.2 Take Profit Level

In order to predict take profit parameter for our model we have to find the predicted value of price change. Take profits is measurement for prediction of trend length which maximize your profit or your prediction without any risk or less risk. We have to find the value of relative impact of any news at time frame't' and compare this relative impact value in old data which have relative impact value equal or nearest to it. After matching the value we will calculate the next time frame price change in the old dataset. This computed price change will be added to open price of the current next time frame and the computated value will be consider as take profit level for the next current time frame in our approach .

$$Take\Pr ofit = Open\Pr ice(t+1) + \Pr edicted\Pr iceChange. \quad (4)$$

3.2.3 Stop Loss

Stop loss level can be any value according to trader risks. It is varies from traders to traders but it is advisable that value should not be greater than take profit.

4. Experimental Result

Here we will compare our proposed model which is hybrid of fundamental trading as well as technical trading with other two technical strategies based on moving average indicator as benchmark model1 and other is Relative strength index as benchmark model 2.

4.1 Data Collection

We have taken EUR/USD pair data from 1-July-2014 00:00 UTC to 15-May -2015 23:59 UTC because we can observe in figure 3 that contribution of Euro and USD is higher than other world currencies.

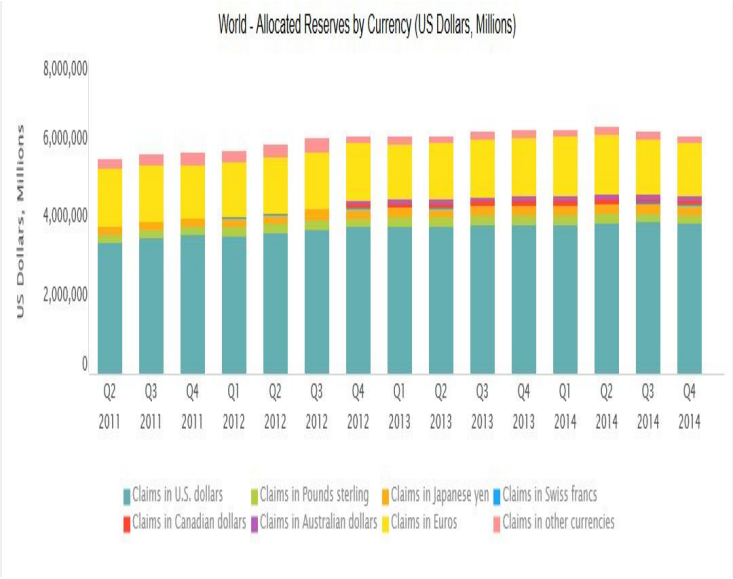

Figure 4 Major currencies reserve by world [4]

4.2 Basic Setting

We will take five period over 180 days in which we have taken chunk of fifteen days for testing our algorithm and comparing it with other two benchmarks in the same parametric conditions. . Initial capital for trading will be 50000$ and at end of every period we will calculate profits and losses on each model.

Profit/loss=Final Revenue- Initial Capital

$$\Pr ofit / Loss = Total\Re venu - IntialCapital \quad (5)$$

(a) Moving Average Indicator (Benchmark model 1)

In this model we analyze the market using moving average indicator and take trade decisions automatically and at the end of each period we calculate profit or loss.

(b) RSI Indicator (Benchmark model 2)

In this model we analyze the market using relative strength index indicator and take trade decisions automatically and at the end of each period we calculate profit or loss.

(c) Sentiment score based algorithmic trading (Proposed Model)

As we have earlier described algorithm, In this model we analyze the market according to above proposed model and at the end of each day we calculated profit or loss.

Table 7: Order Summaries

Period	Moving Average Indicator	Relative Strength Index Indicator	Sentiment score based algorithmic trading(Proposed Model)
Period1	-25741	-49548	**10200**
Period2	122109	-793.87	**177850**
Period3	39169.33	47514.95	**67400**
Period4	19042	76090.73	-84600
Period5	-23558	-48595.37	**-2550**
Average	4711.6	4933.488	**33660**

Table 8: Accuracy Summaries

Period	Moving Average Indicator	Relative Strength Index Indicator	Sentiment score based algorithmic trading(Proposed Model)
Period1	33.3	0.09	**50**
Period2	80	40	**100**
Period3	54	46	50
Period4	42.9	50	**66.6**
Period5	33.3	27.77	**80**
Average	48.7	32.77	**69.32**

Simulation Results: Table no 7 indicates that proposed model (Sentiment score based algorithmic trading) is giving higher returns than other two benchmark (moving average & RSI) and Table no 8 indicates that proposed model is showing 69.32 % average accuracy where other models like moving average is giving the accuracy of 48.7% and RSI is giving accuracy of 32.77 % which is higher than other two model . It indicate that propose model is less risky and more profitable than other two models.

5. Conclusions and Future Work

We have find the accuracy of 69.32 which is higher to the other benchmarks this signifies that our proposed model is predicting the value precisely than other existing model. Along with all these, there are some silent features of our model

1. Our model is using very less data for computation.
2. We are considering all news for the prediction rather than depending on single news for prediction.
3. Our combined approach of fundamental trading and technical trading is also removing the major drawback of technical trading as well as fundamental trading.

Along with features we have also some limitations. In the bubble situation, our model is not allowing to trade. But in future we will try overcome this limitation.

Time	Open	Close	News Time	Currency	News	Relative Impact	Price change
2015.03.01 20:00:00	1.11808	1.11607				1000	-20.1
2015.03.02 00:00:00	1.11609	1.11759					15
2015.03.02 04:00:00	1.11758	1.11787					2.9
2015.03.02 08:00:00	1.11786	1.12244	10:00am	EUR	CPI Flash Estimate y/y	1020	45.8
2015.03.02 12:00:00	1.12244	1.11938					-30.6
2015.03.02 16:00:00	1.1194	1.1186					-8
2015.03.02 20:00:00	1.11862	1.11766					-9.6
2015.03.03 00:00:00	1.11765	1.11886					12.1
2015.03.03 04:00:00	1.11888	1.1205					16.2
2015.03.03 08:00:00	1.12045	1.11584	8:00am	EUR	Spanish Unemployment Chan	1000	-46.1
2015.03.03 12:00:00	1.11583	1.11846					26.3
2015.03.03 16:00:00	1.11847	1.11776					-7.1
2015.03.03 20:00:00	1.11776	1.11827					5.1
2015.03.04 00:00:00	1.11826	1.11738					-8.8
2015.03.04 04:00:00	1.11737	1.11762					2.5
2015.03.04 08:00:00	1.11762	1.11277					-48.5
2015.03.04 12:00:00	1.11277	1.1069	1:15pm	USD	ADP Non-Farm Employment (	970	-58.7
2015.03.04 16:00:00	1.10691	1.10738					4.7

Figure 3 Datasets snapshots of our proposed model [3]

References

[1] Xiaodong Li, Haoran Xie,Ran Wang,Yi Cai,Jingjing Cao,Feng Wang,Huaqing Min,Xiaotie Deng, Empirical analysis: stock market prediction via extreme learning machine, London: Springer-Verlag, 2014.

[2] Arne Thorvald Gierløff Hollum, Borre P. Mosch, and Zoltán Szlávik, Economic Sentiment: Text-Based Prediction of Stock Price Movements with Machine Learning and WordNet , Berlin Heidelberg : Springer-Verlag,2013.

[3]http://thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/news-analytics-product-brochure--oct-2010.pdf.

[4] Currency Composition of Official Foreign Exchange Reserves (COFER), International Financial Statistics (IFS),” http://data.imf.org/”.

[5] http://www.dukascopy.com

[6] EleftheriosSoulas, Dennis Shasha “Online Machine Learning AlgorithmsFor Currency Exchange Prediction” TR-2013-953

[7] Megan Potoski” Predicting Gold Prices” CS229,2013

[8] Chenxu Shao, Zheming Zheng “Algorithmic trading using machine learning techniques”2013

[9] Gabriel Pui Cheong Fung, Jeffrey Xu Yu, and Wai Lam “News Sensitive Stock Trend Prediction”.

[10] Jinli Hu, Amos Storkey “Multi-period Trading Prediction Markets with Connections to MachineLearing.