

Credibility of Social Media Postings: A Genetic Algorithmic Approach to Stock Market Contexts

Amitrajit Chakrabarti

Computer Science and Engineering Department
Jadavpur University, Kolkata 700032, India
Email: amitrajit1@gmail.com

Soumalya Seal

Computer Science and Engineering Department
Jadavpur University, Kolkata 700032, India
Email: soumalya.seal@gmail.com

Uttam K Sarkar

Indian Institute of Management Calcutta
Joka, Diamond Harbour Road, Kolkata 700104, India
Email: uttam@iimcal.ac.in

Abstract— Affordable access to electronic news and social media have increased the propensity of people to browse abundant opinions expressed by others and get influenced by those opinions while taking related decisions. The degree of the uncertainty looming over the optimality of the decision and its associated stake influence the intensity of this inclination. The stock market is one example where the uncertainty is high and so are the stakes. Ordinary investors skim through freely available expert opinions and recommendations in the social media on *buying* or *selling* a stock without knowing how much those advices were worth. Interpretation and assessment of an opinion get complicated because it is expressed in a natural language, such as English, which is not easily amenable to an unambiguous quantification of the expressed opinion. This research proposes a novel method of quantifying unstructured textual opinions of stock market experts in a genetic algorithmic framework. It explores to what extent the stock price movements of some stocks are more in sync with expert recommendations compared to other stocks, and how contrasting the predictions induced by the recommendations of different experts are. Empirical studies have been performed with a large volume of publicly available stock market data and associated expert opinions expressed in various social media. The findings indicate the proposed method to be a credible way of treating opinions in the domain of stock markets. By using the method an investor can empower herself while treating social media information in accordance with its merit.

Keywords— *Opinion mining, Social media, Genetic algorithm, Correlation coefficient, Web-crawler.*

I. Introduction

The complexity of the nature of price movement of stocks in a stock market makes it one of the most difficult prediction problems [Gerasimo 2005, Yang 2006, Philip 2007, Roh 2007]. It's known that a stock market price movement cannot be predicted with high accuracy let alone exactly as that would lead to a collapse of the market. The price movements are rather like a random walk phenomenon [Cootner 1964, Fama 1965, Malkiel 1973]. However, the stake associated with predictions is very high and the user of one model may benefit at the cost of those players in the market whose decisions are based on different perhaps worse models. Consequently, a never-ending evolution of predictive models in this domain continues to flourish. Due to intrinsic intractability of the underlying problem deterministic algorithms don't work well and meta-heuristic methods, such as artificial neural networks and genetic algorithms, are often used[Wu 2001, Smith 2000, Abu 2001].

Experts who give their recommendations on stock transactions in social media presumably use their world knowledge and findings from predictive models while giving an advice or recommendation regarding a stock transaction. Because of the difficulty of making a credible prediction, opinions are often not communicated in very decisive terms by using a single word, that is, BUY, HOLD, or SELL that matters to an investor. Though

such concrete recommendations also exist, the prevailing practice is to express such opinions in a natural language with or without reasons to convince the reader. For example, a rather detailed one as in “*PNB had highest RoA, RoE among PSU banks in FY08-10 but has taken a huge knock as deteriorating asset quality led to elevated provisioning and loss of NII. We believe return ratios are unlikely to improve in the near term on provision for stressed assets, impacting PAT. PNB will be a major beneficiary of MTM reversal on investment book (AFS book proportion of 29%). We recommend BUY from a long term view and maintain our target price at Rs 1156 based on 1.4x revised ABV of Rs 854*” or, a brief opinion as in “*Tata Motors DVR is a SELL with a stop at 290 and look for targets of around 272*”, etc. To ease the interpretability of the textual opinions experts use several keywords whose perceived meanings are expected to be as free from ambiguity as possible but, in reality those appear quite confusing to an average reader. The list of such keywords is long and typical examples include *buy, sell, hold, strong-buy, long-term buy, top pick, underperform, accumulate*, etc.

It is very difficult to extract the intended advice of the expert from the text from a natural language as an expression is not necessarily free from the inherent ambiguity of the underlying vocabulary. For example, whether a “*good buy*” is to be quantified as 4.2 or 4.5 in a [0 to 5] points scale, where 5 indicates strongest buy and 0 the strongest sell recommendation, is not known. Worse, usage of such fuzzy terms is not consistent across expert opinions.

The existence of inconsistency in usage of keywords paves ways for research where human judgment in interpreting unstructured opinion can be fine-tuned based on historical data in a supervised learning framework. While there are hundreds of predictive models for the stock market, not much work is known in the literature on how the keywords used in expert opinions can be suitably quantified despite possible inconsistencies in the available opinions. Penetration of social media has been so deep in shaping our everyday decisions that recent studies indicate even Twitter postings, Google search keywords, and Wikipedia search keywords serve as early indicators of impending stock market transactions. An appropriate quantification of the keyword strings found in opinions expressed in social media would serve as a decision making aid in stock market transactions.

This paper proposes a method of quantifying keywords commonly found in expert opinions using a genetic algorithm. Based on the results the textual opinions of experts can be quantified and interpreted and the impact of consequent decisions to buy or sell can be compared with the true price movements to capture the usefulness of opinions. The observed variations in performance can be used to segregate experts into three categories, namely, (a) those whose recommendations proved correct (b) those for whom no significant link is observed, and (c) those whose opinions were rather misleading. The nature of correlation between recommendations and market performance for a specific stock is also studied based on collective recommendation of different experts in order to explore the variation of predictability across stocks as some stock movement may be relatively more difficult to predict compared to that of others.

This research works with assumption that majority of the experts who offer advice are consistent in their use of keywords related to their advice. Under this assumption the keywords they use should be translatable to numeric values within reasonable error margins. This task entails a complex optimization in assigning values to keywords in order to reconcile past recommendations and actual price movements in the reality. This optimization, in the framework of genetic algorithm, is the core focus of this research.

Over 300 commonly used keywords which are most frequently used by the experts in their advices have been identified. A total of 73 Indian companies whose stocks often surface in experts’ advices have been chosen. These include a variety of sectors including pharmaceutical industry, IT industry, banks, and fast moving consumer goods (FMCG) companies. Stock market data of the Indian market from 2004 to 2013 were considered in the study. This period also covered a boom time and also a low growth period. However, the approach and the proposed method would be applicable to any other stock market context.

The rest of the paper is organized as follows. Section II describes the data and the data collection detail. Section III presents the proposed method. Section IV discusses the empirical findings. Section V draws concluding remarks.

II. Data Capture

Data Collection

Stock price of a company in the market is expected to be governed by investor sentiment, cash flow, marketing policies, debt repayment capability, past performance, nature of customer base, and quality of managers among dozens of other attributes. For a meaningful analysis expert advice data need to be collected in an organized form over prolonged period of time. Moneycontrol.com, one of several websites which contains such data, was considered as a major data source for this research. This website contains advices for a variety of companies and by a number of experts. Most of the advices of moneycontrol.com are also tweeted from their official Twitter account, and the experts also express their views on television channel CNBC 18. Consequently, it has a widespread public reach.

Stock prices of the companies included in the study as listed in the National Stock Exchange were obtained from Yahoo Finance. Three different time frames, 180 calendar days (for short-term investors), 365 calendar days (for mid-term investors) and 730 calendar days (for long-term investors) are considered in the study. Closing stock prices were stored over the entire time period of study so that short-term, mid-term, and long-term return from a stock can be computed with respect to an investment decision at any point of time.

A web-crawler is a software application capable of visiting hyperlinks and extracting data from the associated web pages. A web-crawler was implemented for collecting expert opinions expressed in the internet on the desired dates.

Based on past price movements the returns were computed within designated time intervals of transactions. These rate of returns were normalized to the range 1 to 5 for ease of readability. The smallest (perhaps negative) rate of return was assigned a value of 1 (indicating a recommended SELL), and the highest value as 5 (recommended BUY), and a linear interpolation between 1 and 5 was done for other returns, where 3 corresponds to a neutral outlook (neither BUY nor SELL, that is, a HOLD recommendation for the stock).

Data used in this research included all trading days over nearly nine year period from 22 November 2004 to 12 August 2013. Not all calendar days are trading days. The trading days, identified by the index i below, are 1,2,3, ..., where $i=1$ corresponds to the first trading day, that is, 22 November 2004, and the highest value of i corresponds to the last trading day, that is, 12 August 2013.

A total of 73 companies were considered in the study and are identified by the index $j=1,2,\dots,73$ below. The companies were selected from diverse sectors including FMCG (fast moving consumer goods), banks, telecom, automobiles, IT, and pharmaceuticals. Opinions of 25 different experts, indexed by $k=1,2,\dots,25$, were analyzed in this paper. For maintaining privacy of the company and the individual experts whose opinions have been mined, the actual names have not been mentioned in the paper and the indices have been used instead. A recommendation is made for a time duration in trading days (denoted by d below). However, fuzzy durations such as *short run*, *long run* are more common and a value of d is then chosen based on popularly perceived meaning of such fuzzy durations.

For all trading days in the time range under study the stock prices are obtained for the specified set of stocks. A set of experts who offered advice and recommendations in plain English for these stocks over these trading days was identified. A set of three hundred commonly occurring keyword strings in expert opinions was identified. Whenever any of these experts expressed opinions on any of these stocks, the textual opinion was captured and parsing was done to check the presence of one or more these keywords strings (such as *good trading bet*, *may see more pain*, *jump*, *not much upside in*, *profit booking*, *may go up*, *can appreciate*, *may not fall*, *good stock*, *looks promising*, *started weakening*).

In what follows, a company and its associated stock have been used synonymously. The collected data and associated processing have been described using the following notation and definitions:

$V_{i,j}$: Closing price of the j -th stock on the i -th trading day; $i=1,2,\dots,T$, where T is the last trading of the calendar time interval $[T_1, T_2]$; in this study, $T_1=22$ November 2004 and $T_2=12$ August 2013; $j=1,2,\dots,L$, where L is the number of companies (or stocks); in this study $L=73$.

$A_{i,j,d,e}$: Advice or recommendation (text string) of expert e for stock j made on the i -th trading day for a period of d trading days.

K_i : The i -th keyword string found in popular recommendations that are considered in this study; $i=1,2,\dots,m$; in this study $m=300$

w_i : Weight assigned to K_i to quantify the valuation of K_i , $1 \leq w_i \leq 5$; 5 indicates strongest *BUY* and 1 indicates strongest *SELL* sentiment of a recommendation. An m -vector of pairs $\langle (K_1, w_1), (K_2, w_2), (K_3, w_3), (K_4, w_4), \dots, (K_m, w_m) \rangle$ will be termed a *complete quantification* of all keywords.

$Y_{i,j,d,e,r}$: =1, if $A_{i,j,d,k}$ contains K_r ,
=0, otherwise. Opinions may not be available for all possible combinations of i,j,d , and e , and only the available recommendations that contains a keyword string of interest, captured by Y , need to be processed.

$Score(A_{i,j,d,e}) = (\sum_r Y_{i,j,d,e,r} w_r) / \sum_r w_r$. That is, *Score* of a recommendation is the average of the weights of all keyword strings found in the recommendation.

$R_{i,j,d}$: Return from the j -th stock based on a transaction decision on trading day i for a period of d trading days, that is, till the trading day $i+d$. Typically, logarithmic returns are used in the literature. In this work, $R_{ij,d} = \log(V_{i,j,i+d} / V_{i,j,d})$ for a buy or hold decision and $R_{ij,d} = \log(V_{i,j,d} / V_{i,j,i+d})$ for a sell decision. Returns are typically computed for some specific values of d such as daily return ($d=1$), weekly return ($d=5$, wherever there are five trading days in a week) etc.

III. Proposed Genetic Algorithmic Framework

Genetic Algorithm (GA) [Holland 1992, Goldberg 1989, Forrest 1993] is one of the meta-heuristic techniques for solving computationally hard optimization problems. The method attempts to mimic natural selection, crossover, and mutation repeatedly over generations of a population. With time the chances of survival of members with above average physical fitness increase. Consequently, the average fitness improves over generations. Applied to the context of optimization, feasible solutions based on some greedy heuristics (even arbitrary feasible solutions at times) play the role of initial population. Fitness of each feasible solution is computed as the associated value of the objective function. Two feasible solutions are then combined in an effort to create another feasible solution that retains good parts of both of them while avoiding their weaker features and is thus expected to have higher fitness than that of either of them. The process is repeated till a specified stopping criterion is satisfied. The fittest *solution* generated in the process is considered to be the output of the method.

The pseudo code of a GA looks like:

Start

Choose initial population of suitable size N;

Choose a fitness function for computing each member's fitness;

Repeat //Iterate over several generations

Repeat //Each member of a generation dies leaving the next generation behind

Select two parents where the chance of selection is proportion to one's fitness;

Apply crossover operator to generate a child;

Apply mutation operator to incorporate limited mutation, if at all, of the child;

Until N children are generated;

Replace the parent population by the N children to get the next generation;

Compute average fitness of the new generation;

Until Average fitness ceased to improve or specified computational resources got exhausted;

Present the fittest member in the present generation is the solution;

End

The key ingredients in applying genetic algorithm to a problem are (i) definition and encoding of a member of the population (ii) identification of the population size and the initial population (iii) designing an appropriate fitness function (iv) devising the crossover operation for producing the next generation (v) devising the mutation operation and (vi) identifying the stopping condition. These stages are addressed below:

Population Member Definition and Encoding

Keyword strings such as “buy”, “strong sell”, “not worth buying” etc are used by experts in their textual messages while conveying their opinions. The central theme of this research is how each of these keywords needs to be quantified and interpreted.

A member of the population is a vector of m pairs $\langle (K_1, w_1), (K_2, w_2), (K_3, w_3), (K_4, w_4), \dots, (K_m, w_m) \rangle$, where $1 \leq w_i \leq 5$ for $i=1, 2, \dots, m$, where w_i quantifies the sentiment of keyword string K_i . In this study $m=300$. A member represents a specific mapping of each keyword string to a numeric equivalent weight in a 1 to 5 point scale, where 1 indicates *strong sell* and 5 indicates *strong buy* sentiment of the recommender.

Creation of Initial Population

An initial assignment of value w_i is made to K_i based on popular perception of the keyword string K_i for all $i=1, 2, \dots, m$, where $1 \leq w_i \leq 5$. For example, “strongest buy” has been assigned the value 5, “strongest sell” to 1, “buy” to 4, etc. The proposed method is based on the acknowledgment of the fact that a human perception of text strings, though not grossly erroneous, is unlikely to be optimal with respect to the objective of assigning the best possible value to the keywords that explain the relationship between past recommendations and associated price movements.

The size of the population was varied from 500 to 1000. For a particular population size ($=n$, say), the i -th member of the population, denoted as P_i , was derived from the initial assignment $\langle (K_1, w_1), (K_2, w_2), (K_3, w_3), \dots \rangle$.

$(K_4, w_4), \dots, (K_m, w_m)$ by randomizing it as $P_i = \langle (K_1, w_1 + f_{i1}z_1), (K_2, w_2 + f_{i2}z_2), (K_3, w_3 + f_{i3}z_3), (K_4, w_4 + f_{i4}z_4), \dots, (K_m, w_m + f_{im}z_m) \rangle$, where $z_j = \min\{5 - w_j, w_j - 1\}$, and each $f_{ij}, i=1,2, \dots, n, j=1,2, \dots, m$, is a random number drawn from the range $[-0.30, +0.30]$. That is, each weight is randomly altered by a maximum amount of 30% of the smaller of the two values by which it can increase or decrease over a range of $[1,5]$ with respect to its initially assigned value. For example, suppose, for a keyword “buy” the initially assigned value is 4. This value can be increased by a maximum of $(5-4)$, that is 1, and it can be decreased by a maximum amount of $(4-1)$ or 3. Minimum of 1 and 3 is 1, and 30% of 1 is 0.3. So, after randomization, the value generated against 4 would be somewhere in the range $[3.7, 4.3]$. This method ensures that the weights will always lie in the range $[1,5]$. Moreover, initial values which are close to either extreme and thus reflect strong conviction of the initial setting are altered far less compared to those values which are close to the center reflecting neutral and unclear sentiments. The logic being if the keyword string “buy” is popularly perceived to have a value of 4 in a 1 to 5 points scale, it’s likely that the magic number that best explains its usage in reflecting recommendations may be expected to be somewhere in the range $[3.7, 4.3]$. Different percentages have been tried and 30% seemed to capture the uncertainty well.

Fitness Computation

For computing the fitness of a member P_r of population, in the context of recommendations made for d trading days, using the data over a time horizon $[T_1, T_2]$ of t trading days the effect of transaction decisions induced by P_r is first computed. Here, T_1 and T_2 denote calendar days where T_1 corresponds to trading day 1 and T_2 corresponds to trading day t , where $t \geq d$. By sliding $[T_1, T_2]$ these effects are listed. This list of values was then compared with the ideal return over time window $[T_1, T_2]$. The higher the match the higher the fitness. The procedure of computation of fitness is described below:

Begin

For j=1 to L //for all stocks

*Compute AverageScore[j] for Stock j by averaging all values of Score(A_{i,j,d,e}) available over [T₁,T₂]
 in accordance with P_r;*

GrandReturn=0; //Total return from all cycles

End for j;

For i=1 to (t-d) //for all d consecutive trading day cycles in [T₁,T₂]

TotalReturn=0; //Total return from all stocks over a cycle

For j=1 to L do //for each stock whether to SELL, HOLD, or BUY this stock

If AverageScore(j) < 2.5 then Transaction[j]=SELL

Else if AverageScore(j) < 3.5 then Transaction[j]=HOLD

Else Transaction[j]=BUY;

*TotalReturn=TotalReturn + R_{i,j,d}; //Cumulate return over stocks based on Transaction[j]
 //for each stock j*

End for j;

GrandReturn=GrandReturn+TotalReturn;

End for i;

*AverageReturn=GrandReturn/((t-d)*L); //Average return per stock over all iterations*

End;

AverageReturn indicates the average return that would have been realized per stock if all stock transaction decisions were taken based on the average recommendations score of all experts on each stock. Each transaction decision was taken by the following rule based on the value of average recommendation score: 1-2.5: SELL, 2.5-3.5: HOLD, and 3.5-5: BUY.

An *ideal* transaction is defined to be a “BUY” if $V_{i,j,i+d} > V_{i,j,d}$ and a “SELL” if $V_{i,j,i+d} < V_{i,j,d}$. In a similar manner, for the same time time period $[T_1, T_2]$ the *ideal* return (that is, when all transactions were *ideal*) and their average, say, *AverageIdealReturn* is also computed.

So, for each time window $[T_1, T_2]$ of the available dataset from the past a pair (*AverageReturn*, *AverageIdealReturn*) is computed. By varying $[T_1, T_2]$, a list of pairs is constructed. The correlation coefficient

between *AverageReturn* and *AverageIdealReturn* is then computed. It may be noted the higher the correlation coefficient, the higher the closeness between *recommendation-induced* transaction and an *ideal* transaction. So, the correlation coefficient can be taken as a measure of the fitness of P_r leading to the recommendations. Since the correlation coefficient (c) may vary between -1 to +1, the fitness of P_r is computed as $(1+c)/2$ so that the fitness gets standardized into the range $[0,1]$.

In the same method, for any specified P_r the correlation coefficient for a specified stock, or for a specific expert can be computed by restricting to only that stock or only to that expert, respectively.

Crossover Process and Creation of Next Generation

Two members, that is, the two parents, are chosen from the population where the chance of any member getting selected is proportional to its fitness value. Let the two parents selected at any iteration be P_a and P_b , where,

$P_a = \langle (K_1, x_1), (K_2, x_2), (K_3, x_3), (K_4, x_4), \dots, (K_m, x_m) \rangle$ and $P_b = \langle (K_1, y_1), (K_2, y_2), (K_3, y_3), (K_4, y_4), \dots, (K_m, y_m) \rangle$. Let F_a and F_b be the fitness values of P_a and P_b , respectively. From the two parents P_a and P_b , a single child C is generated as

$C = \langle (K_1, z_1), (K_2, z_2), (K_3, z_3), (K_4, z_4), \dots, (K_m, z_m) \rangle$, where $z_i = (x_i + y_i)/2$ if i is odd, and $z_i = F_a x_i + F_b y_i$ if i is even, $i=1,2,\dots,m$.

The process is repeated with the population members until as many child members are generated. The original population is then discarded and the children thus created form the next generation.

Mutation Process

Mutation was kept at 5 percent level. That is, a generated child is mutated with a probability of 0.05. Once a decision is taken to mutate a child one of its m weights is chosen at random and the weight is replaced by a random real number between 1 and 5.

Termination Condition

The process of recreations through crossover and mutation is repeated for 1000 generations. The fittest member of the final generation is identified as the solution.

IV. Experimental Results

Stock market related recommendations are made keeping some investment time horizon in mind. For example, a recommendation made for a period of 3 years should not be assessed by observing the performance of the stock only after 2 months following the investment. However, experts do not precisely specify the time frame for which they make a recommendation. Instead, they use more generic time durations such as “*short-term*”, “*mid-term*”, and “*long-term*”. While there is no universal consensus on the meaning of these *terms*, the essence of this study can be captured if *short term* is assumed to be less than a year, *mid-term* to be more than a short-term but less than two years, and *long-term* to be two years or more. Results presented in this Section have accordingly been divided into three categories – *short-term* (six calendar months), *mid-term* (one calendar year), and *long-term* (two calendar years). Calendar month is mentioned to avoid any confusion with the number of trading days because for any time duration the latter is less than intermediate calendar days because of stock market closures due to weekends and other planned vacations.

Ten companies are chosen for *short term*, *mid-term*, and *long-term* investment horizons for presenting the empirical findings. For each of these three categories of investments the effectiveness of keyword quantification is obtained by computing the correlation coefficient between actual price movements and the price movements induced by the recommendations based on the quantified values of keywords as explained in Section III. Those processing data in the stock market context would know the value of any correlation between predicted transaction decision and ideal transaction decision for making profits is likely to be very close to zero because of the inherent nature of the prediction problem. Any perceptible deviation from zero would be suggestive of the strength of the method.

Opinions of 25 experts who commented on and made their recommendations in the electronic and social media about one or more of the stocks under study were chosen for analysis. Their performances were analyzed for

each of these three categories of investments. The correlation coefficient between the values of the keywords in their advice and the original value obtained by analyzing the stock market trends are calculated and compared. The list of experts is arranged in decreasing order of their correlation coefficients. The expert having the highest correlation coefficient is most likely to give a useful advice for the ten companies which we have shown. The experts and the company names have not been explicitly mentioned to protect their identity. However, it is not at all necessary for the study to name them as the study is concerned only with the relative variations.

The genetic algorithm was used in an effort to capture the optimal weight of each keyword string that reconciles recommendations with the reality. The weights thus identified formed the basis of all further computations related to the assessment of expert opinions. Each textual opinion of an expert made on a particular day about a particular stock keeping a specified time horizon in mind, may be a sentence or even paragraph of plain text, was assigned a recommendation score based on average valuation of all keyword strings in the recommendation and various correlations were computed as explained in Section III.

Short-term Opinion: Table 1 shows the correlation coefficients for the ten companies.

Table 1. Learning of keyword valuation in short term

Company	1	2	3	4	5	6	7	8	9	10
CCoeff(C)	.295	.269	.263	.210	.169	.160	.125	.118	.112	.111

The detailed preparation of data for computing the correlation coefficient for a company (or stock) was explained in Section III. A high value indicates the stock is more amenable to prediction based on expert opinions compared to other stocks with low values of correlation coefficients. Since the number of opinions available was not the same for all stocks the weighted average of these correlation coefficients, where each correlation coefficient value, C, was weighted by the number of instances (M) of opinion available for the stock ($= \sum (MC)/\sum M$) was also computed. The value is 0.176.

It may be noted that a weighted correlation coefficient value of 0.176, can be considered quite away from zero in the stock market prediction context, and can be interpreted to be significantly high. The significance can also be computed following standard statistical procedures. A significant value shows the effectiveness of the proposed method.

It's known that stock market expert opinions are sometimes wrong and misleading. Correlation of the prediction and actual performance is, therefore, of great significance. The proposed method also facilitated computation of these correlation coefficients in order to assess the performance of specific experts. In Table 2 the performance of the experts are shown in the order from the best to the worst. In the Table, SI No indicates the id of the expert and CCoeff indicates the correlation coefficient between returns based on her recommendations and the ideal return for all stocks she opined for. The computation was done as explained in Section III. The results indicate experts clearly belonged to three different categories – (a) Those showing significantly positive, say 0.2 and above, correlation coefficients. Their opinions are certain worth looking into. (b) Those having near zero values. This shows their recommendations are devoid of useful contents for an investor, and (c) Those showing negative correlation coefficients. The experts in the last group are either deficient in their competence or their opinions might have been deliberately biased by other vested interests, not an uncommon phenomenon, and are detrimental to ordinary investors who might take those opinions by face value. Of twenty five experts, seven (28%) belonged to category (a), twelve (48%) belonged to category (b), and the remaining six (24%) were in category (c). This shows a huge majority of the expert opinions are either useless or misleading.

Table 2: Performance of Experts

SI No	CCoeff	SI No	CCoeff	SI No	CCoeff	SI No	CCoeff	SI No	CCoeff
-------	--------	-------	--------	-------	--------	-------	--------	-------	--------

1	.648	6	.317	11	0	16	0	21	-.368
2	.621	7	.208	12	0	17	0	22	-.388
3	.601	8	.097	13	0	18	-.093	23	-.395
4	.430	9	.084	14	0	19	-.168	24	-.486
5	.384	10	.065	15	0	20	-.296	25	-.597

Medium-term Opinion:

The same experiment, as done in case of short –term, was carried out for mid-term opinions as well. Table 3, when compared with Table 1, indicates that for top ten companies medium term recommendations were marginally better compared to short-term predictions.

Table 3. Learning of keyword valuation in medium term

Company	1	2	3	4	5	6	7	8	9	10
CCoeff(C)	.344	.262	.251	.240	.239	.199	.192	.160	.127	.111

Weighted average Correlation Coefficient = $\sum (MC)/\sum M = 0.193$. The performance of experts in mid-term prediction is shown in Table 4.

Table 4: Performance of Experts in medium term

Sl No	CCoeff	Sl No	CCoeff	Sl No	CCoeff	Sl No	CCoeff	Sl No	CCoeff
1	.692	6	.300	11	.108	16	0	21	-.076
2	.474	7	.237	12	.105	17	0	22	-.113
3	.421	8	.162	13	.100	18	0	23	-.297
4	.375	9	.150	14	.024	19	-.011	24	-.303
5	.324	10	.113	15	.009	20	-.013	25	-.433

Table 4 shows, of twenty five experts, seven (28%) belonged to category (a), fifteen (60%) belonged to category (b), and the remaining three (12%) were in category (c). When compared with Table 2, it indicates the percentage of useful opinions more or less remained the same but a relatively fewer percentage of experts gave misleading opinions.

Long-term Opinion: Table 5 and 6 below show the results for long-term time horizons.

Comparison of Table 5 with Table 1 and Table 3 shows in the long-term the quality of opinions is superior to those of short term. This is expected because in the stock prices in the long-term better reflect the fundamentals of the company and depend less on speculative elements.

Table 5. Learning of keyword valuation in long term

Company	1	2	3	4	5	6	7	8	9	10
CCoeff(C)	.548	.489	.344	.263	.234	.216	.193	.189	.184	.131

Weighted average correlation coefficient = $\sum (MC)/\sum M = 0.249$

Table 6: Performance of experts in long term

SI No	CCoeff	SI No	CCoeff	SI No	CCoeff	SI No	CCoeff	SI No	CCoeff
1	.810	6	.182	11	0	16	0	21	-.328
2	.795	7	.182	12	0	17	0	22	-.419
3	.434	8	.143	13	0	18	0	23	-.449
4	.256	9	.137	14	0	19	0	24	-.561
5	.255	10	.118	15	0	20	-.259	25	-.669

V. Conclusion

The unpredictable volatility of the stock market and consequent effort to model it in order to reap above-average returns from the market continues to remain a coveted but illusory goal. The efficient market hypothesis (EMH) rules out any prospect of sustainable superior returns in an ideal market where all investors are well informed. In reality, investors take decisions based on non-homogeneous and disparate information. Social, political, and speculative behavior of powerful brokers also distort the market. The end result is that the stock market exhibits more idiosyncrasies than what simple models can capture and market prices often exhibit existence of prospects for superior returns, especially in the short term.

The uncertainty in the stock market, the difficulty of building a predictive model from the noisy data, and the prospect of making money from market foresight attract the interests of ordinary investors toward expert opinions on the subject. While the investors are not too inclined to take such opinions by face value, they do not ignore these either and face considerable difficulty in deciding how much importance the opinions deserved. Overabundance of information freely available in the social media, which are often more inconsistent than not, could be a source of strength as well as of weakness in the decision making capability of individuals in the modern information-dominated society. Objective quantification of subjective opinions becomes important in this context.

This paper proposed a methodology to quantify subjective opinions in a manner the performance of those who opine in the stock market context can be assessed in an objective way based on historical data. The underlying analytics facilitates ranking companies with respect to the predictability of their performance as well as ranking experts with regard to the quality of their recommendations. The results show when initial human judgment regarding quantitative semantics of natural language keywords strings can be fine-tuned with the help of past data in a supervised learning framework using a genetic algorithm, it is possible to improve the quantification in a significant way. Findings indicate the proposed approach consistently demonstrated its capability in short-term, medium-term, and long-term investment horizons. These findings can be used in portfolio management applications as well as individual transaction decisions whenever expert opinions available in social media are considered in the decision process.

A more sophisticated semantic analysis to find the recommendation score of a text opinion is expected to improve the results further. Moreover, if an appropriate outlier analysis on expert opinion data and company performance data is first carried out and the filtered data are used in the proposed model it would be expected to model the reality better.

References

- Abu-Mostafa Y. S., Atiya A. F., Magdon-Ismael M., and White H. 2001, "Neural Networks in Financial Engineering", IEEE Transactions on Neural Networks, vol. 4, no. 12, pp. 653-656.
- Cootner P. H. 1964, "The random character of stock market prices", MIT Press. ISBN 978-0-262-03009-0.
- Fama E. F. 1965, "Random Walks In Stock Market Prices". *Financial Analysts Journal* 21 (5): 55-59.

- Forrest S. 1993, "Genetic algorithms: Principles of natural selection applied to computation", *Science* 261: pp. 872-878
- Gerasimo M., Konstantions P., Yannis T., and Babis T. 2005, "Intelligent Stock Market Assistant using Temporal Data Mining", In Proc. of 10th Panhellenics Conference in Informatics (PCI05), Greece.
- Goldberg, D. E. 1989, "Genetic Algorithms in Search, Optimization, and Machine Learning", Reading, MA: Addison-Wesley, ISBN 0201157675.
- Holland J. H. 1992, "Genetic algorithms", *Scientific American*, July pp. 66-72.
- Malkiel B. G. 1973, *A Random Walk Down Wall Street* (6th ed.). W.W. Norton & Company, Inc. ISBN 0-393-06245-7.
- Philip M. T., Paul K., Choy S. O., Reggie K., Ng S. C., Mak J., Jonathan T., Kai K., and Tak-Lam W. 2007, "Design and Implementation of NN5 for Hong Stock Price Forecasting", *Journal of Engineering Applications of Artificial Intelligence*, vol. 20, pp. 453-461.
- Roh T. H. 2007, "Forecasting the Volatility of Stock Price Index", *Journal of Expert Systems with Applications*, vol. 33, pp. 916-922.
- Smith K. A. and Gupta J. N. 2000, "Neural Networks in Business: Techniques and Applications for the Operations Researcher", *Computers & Operations Research*, no. 27, pp. 1023-1044.
- Wu X., Fund M., and Flitman A. 2001, "Forecasting Stock Performance using Intelligent Hybrid Systems", *Springerlink*, pp. 447-456.
- Yang Q. and Wu X. 2006, "10 Challenge Problems in Data Mining Research", *International Journal of Information Technology & Decision Making*, vol. 5, No. 4, pp. 1-8.

[Amitrajit Chakrabarti, Soumalya Seal, and Uttam K Sarkar] © 2014. The authors assign to ACIS and educational and non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to ACIS to publish this document in full in the Conference Papers and Proceedings. Those documents may be published on the World Wide Web, CD-ROM, in printed form, and on mirror sites on the World Wide Web. Any other usage is prohibited without the express permission of the authors.