StockAssIstant: A Stock AI Assistant for Reliability Modeling of Stock Comments

Chen Zhang
360 Search Lab
Beijing, China
zhangchen010295@163.com

Yijun Wang
LineZone Data
Hangzhou, China
wangyijun@linezonedata.com

Can Chen
LineZone Data
Hangzhou, China
chencan@linezonedata.com

Changying Du
360 Search Lab
Beijing, China
ducyatict@gmail.com

Hongzhi Yin*
The University of Queensland
Brisbane, QLD, Australia
h.yin1@uq.edu.au

Hao Wang†
360 Search Lab
Beijing, China
cashenry@126.com

ABSTRACT

Stock comments from analysts contain important consulting information for investors to foresee stock volatility and market trends. Existing studies on stock comments usually focused on capturing coarse-grained opinion polarities or understanding market fundamentals. However, investors are often overwhelmed and confused by massive comments with huge noises and ambiguous opinions. Therefore, it is an emerging need to have a fine-grained stock comment analysis tool to identify more reliable stock comments. To this end, this paper provides a solution called *StockAssIstant* for modeling the reliability of stock comments by considering multiple factors, such as stock price trends, comment content, and the performances of analysts, in a holistic manner. Specifically, we first analyze the pattern of analysts' opinion dynamics from historical comments. Then, we extract key features from the time-series constructed by using the semantic information in comment text, stock prices and the historical behaviors of analysts. Based on these features, we propose an ensemble learning based approach for measuring the reliability of comments. Finally, we conduct extensive experiments and provide a trading simulation on real-world stock data. The experimental results and the profit achieved by the simulated trading in 12-month period clearly validate the effectiveness of our approach for modeling the reliability of stock comments.

KEYWORDS

Stock comment, Time-series, Reliability modeling.

ACM Reference Format:
Chen Zhang, Yijun Wang, Can Chen, Changying Du, Hongzhi Yin, and Hao Wang. 2018. *StockAssIstant*: A Stock AI Assistant for Reliability Modeling of Stock Comments. In *KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, August 19–23, 2018, London, United Kingdom.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3219819.3219964

*Corresponding Author.
†Corresponding Author. Note that the intellectual property rights of this paper belong to 360 Search Lab.

KDD '18, August 19–23, 2018, London, United Kingdom

ACM ISBN 978-1-4503-5552-0/18/08…$15.00
https://doi.org/10.1145/3219819.3219964

1 INTRODUCTION

An increasing number of stock comments are becoming available at stock forums, such as Yahoo! Finance message board [1] and Sina Financial Planner[2]. A stock comment refers to the analysis and prospect of a specific stock movement by an analyst [3]. Indeed, stock comments have been a rich source of information for investors to understand market trends or stock volatility [1, 11]. For instance, researchers have made some efforts to investigate the relationship between stock comments posting activities and stock trading activities [2, 11, 17, 24, 44]. These studies usually focused on capturing coarse-grained opinion polarities or understanding the major impact of comments on market movements [12]. Moreover, some other works [1, 19, 23, 28] used classification models such as SVMs to explicitly predict stock trends by exploiting sentiment features from stock comments.

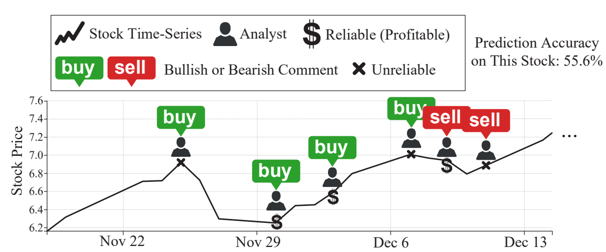


Figure 1: A comment sequence from an analyst.

However, to reliably predict stock trends is challenging, as will be demonstrated in Figure 4, owing to the fact that the stock prices are affected by many uncertain economic-political factors, such as company performances, economic climate, government policies, and even bombshells of the companies. Therefore, due to the noisy

[1] http://finance.yahoo.com/
[2] A professional stock analysis forum. http://licaishi.sina.com.cn/.
[3] Stock comments can also focus on whole market or a sector. In this paper, we only consider stock comments on individual stocks.

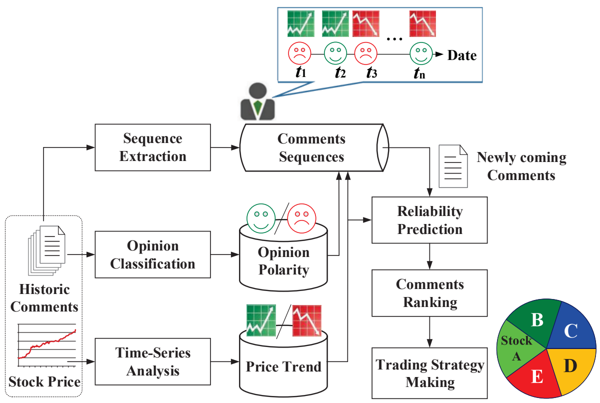


Figure 2: The framework of our approach.

and bias nature of stock comments, the reliability of stock comments is of greatest concern to investors. There is an urgent need to have a fine-grained comment analysis tool to identify more reliable comments, which can help investors to better understand the market status and guide them to build profitable trading strategies. Nonetheless, it is nontrivial to determine if a comment is reliable, especially in the real-world market where investors are overwhelmed by massive comments with conflict or ambiguous opinions.

As shown in Figure 1, to measure the reliability of stock comments, one should consider several factors, such as historical stock prices, current market status and the opinion polarities hidden in comment text. Also, the historical performances of analysts, i.e. their prediction accuracies and dynamic opinion shifting behaviors, are also key factors. In fact, the opinions of an analyst are evolving along with dynamic market situations. They may hold the same opinion or shift to the opposite direction. For example, some analysts are more likely to change their opinions about the market trend if their previous judgments are not consistent with current market situations. Thus, the reliability of stock comments should be modeled in a time-evolving manner. However, previous studies have limited efforts on capturing the opinion shifting patterns of analysts. As a result, how to model the opinion dynamics and incorporate all of the above factors in a unified framework remains to be a challenge.

To this end, we first analyze the historical stock comments in a time-evolving manner, and uncover some important phenomena regarding the coherence characteristics of analysts' opinions and the opinion shifting patterns. Then, we provide a systematic approach to model the reliability of each stock comment. Figure 2 shows the framework of the approach. Specifically, we employ the Factorization Machines (FM) to detect the opinion polarity of each stock comment. Next, we extract a number of key features from the time-series data constructed by using the semantic information of stock comment text, stock prices and analyst behaviors. Lastly, we present an ensemble learning based approach which can exploit the strengths of classification model and Time-Series Analysis (TSA) model, i.e. Support Vector Machines (SVMs) and AutoRegressive Moving Average (ARMA), to measure the reliability of each stock comment. With the identified reliable comments, we

Figure 3: A stock comment from an analyst.

can build profitable trading strategies. To validate the effectiveness of our approach, we perform extensive experiments on real-world data and provide a stock trading simulation. The experimental results and the profit achieved by the simulated trading in 12-month period show that our approach is effective in modeling the reliability of stock comments and can be applicable for various finance related applications, such as financial trend prediction, portfolio management and automated trading.

To the best of our knowledge, this is the first attempt to provide a fine-grained analysis of the reliability of stock comments. The main contributions of this paper are as follows:

(1) *The discovery of the coherence characteristics of analysts' opinions and the opinion shifting patterns.*
(2) *Multi-facet feature selection from stock comment text, stock prices, and analyst behaviors.*
(3) *An ensemble learning based framework to model the reliability of stock comments.*
(4) *The development of effective stock trading strategies based on reliability modeling of stock comments.*

Overview. The rest of this paper is organized as follows: Section 2 describes the overview of the used data. In Section 3, we provide a comprehensive observation about characteristics of the stock analysts and give the corresponding explanations of some meaningful phenomena. Then, Section 4 presents a hybrid stock comments recommendation approach. Experiments and a market trading simulation are performed in Section 5. Section 6 introduces the related work. Finally, the conclusion is summarized in Section 7.

2 DATA DESCRIPTION

Stock Comment Data. Figure 3 shows an example from Yahoo! Finance message board [1]. It indicates that an online analyst called *allan* expressed the *bullish* opinion *8 days ago* on the stock with the symbol of *IBM*, and gave his reason: *there is a support*. Therefore, a stock comment usually contains five elements: the analyst, the stock symbol, the opinion polarity, the comment text and the posted timestamp.

Here we formally represent the stock comments dataset as $C = \{c_1, c_2, \dots, c_{|C|}\}$, the set of stock symbols as $S = \{s_1, s_2, \dots, s_{|S|}\}$, and the set of analysts as $A = \{a_1, a_2, \dots, a_{|A|}\}$. Please note that $|\cdot|$ denotes the size of the set. Based on that, we give the definition for each comment.

Definition 2.1. (***comment unit***) *A comment unit is defined as a sextuple* $c_i = \{d^{(c_i)}, a^{(c_i)}, s^{(c_i)}, t^{(c_i)}, o^{(c_i)}, r^{(c_i)}\}$, *which contains the following information:*

- $d^{(c_i)}$: *the comment text;*

- $a^{(c_i)}$: *the analyst who posts the comment;*
- $s^{(c_i)}$: *the stock symbol mentioned in the comment;*
- $t^{(c_i)}$: *the timestamp when the comment is posted;*
- $o^{(c_i)}$: *the opinion polarity held by the analyst; and*
- $r^{(c_i)}$: *the label for reliability.*

Here $o^{(c_i)}$ and $r^{(c_i)}$ are two unknown variables that we need to predict. Besides, $o^{(c_i)}$ is a boolean variable, either bullish (1) or bearish (−1); $r^{(c_i)}$ is also a boolean variable, either reliable (1) or unreliable (−1). When the target stock's price increases in the next day and the opinion polarity is bullish (or bearish), $r^{(c_i)}$ is set to 1 (or −1). Otherwise, when the price decreases and the opinion polarity is bullish (or bearish), $r^{(c_i)}$ is set to −1 (or 1).

Note that we choose the stock trend in the next day to assess if a comment is reliable, because for both human (analysts) and algorithms, the performance of short term (e.g. daily) prediction is usually better than long term (e.g. weekly or monthly) prediction [14, 42]. In fact, most comments do not explicitly indicate the width of prediction window, e.g. the granularity of one day/week/month. Furthermore, a comment may even make multiple predictions on various windows. Thus, to quantitatively determine the granularity of a comment, i.e. if it is relevant to short term or long term prediction, is rather difficult (even for human). So, as the first attempt to model the reliability of stock comments, we currently do not consider the granularity and simply treat it as the short-term prediction. However, future work will focus on identifying the comment granularity and use it to perform more accurate modeling.

Based on the definition of *comment unit*, we give another definition: *comment sequence*. Figure 1 shows an example.

Definition 2.2. (***comment sequence***) *A comment sequence $l(i,j)$ is a time-ordered sequence of comment units, which is denoted by $c_1 \rightarrow c_2 \rightarrow \ldots \rightarrow c_{|l(i,j)|}$, where $|l(i,j)|$ denotes the length of $l(i,j)$. In $l(i,j)$, every comment unit was posted by the same analyst a_i on the same stock s_j.*

We collect 187,782 comments posted by 1,154 analysts from Sina Financial Planner[2], in which 2,969 stocks were analyzed. The time span is from August 2014 to December 2016. We convert the raw comments into *comment sequences* and remove the *comment sequences* with the length less than 5 to avoid the data sparsity problem. The *comment units* holding neutral opinions are also filtered as they may not be useful for investors to make trading decisions.

Stock-related Data. The stock price data is the adjusted closing price. The stock sector data contains ten sectors based on the industry fields of firms. As shown in Table 1, those sectors are *Consumer Discretionary, Industrials, Information Technology, Energy, Utilities, Materials, Telecommunication Services, Health Care, Consumer Staples* and *Financials*. These data are downloaded from RiceQuant[4].

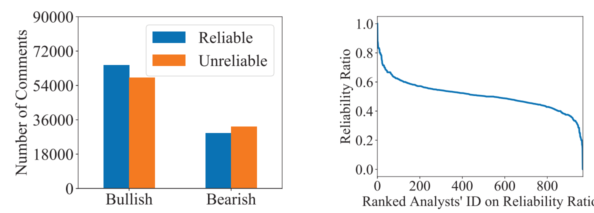
3 OBSERVATION OF ANALYST BEHAVIORS

In this section, we analyze the *comment sequences* in a time-evolving view, and observe some interesting phenomena regarding the coherence characteristics of analysts' opinions and the opinion shifting patterns. To the best of our knowledge, such characteristics and patterns have not been investigated in previous research. However, we claim that they are potentially useful for extracting the analyst-related features in Table 2.

[4]A Chinese stock trade platform, https://www.ricequant.com/.

Table 1: Sectors of stock symbols

Category	# Covered Symbols
ConsumerDiscretionary	129
Industrials	207
InformationTechnology	154
Energy	19
Utilities	36
Materials	131
TelecommunicationServices	5
HealthCare	38
ConsumerStaples	48
Financials	84



(a) Comments (b) Analysts

Figure 4: Distributions of opinion and reliability.

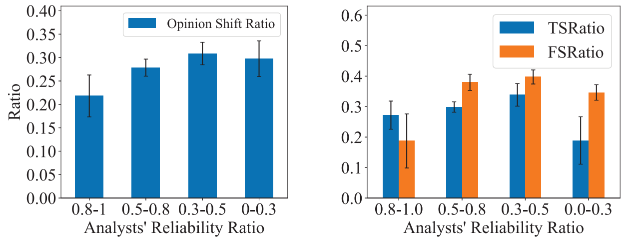
Before the investigation, we list three questions which may help to explain the relations between the opinions expressed by analysts and the reliability of stock comments posted by analysts:

- *Which opinion polarities do analysts tend to express in a comment; and to what extent are these polarities reliable?*
- *For one stock, do analysts tend to hold coherent opinions or to shift their opinions frequently?*
- *Under what circumstances do they tend to hold/shift their opinions; and is to hold/shift opinions a good strategy?*

3.1 Distributions of Opinion and Reliability

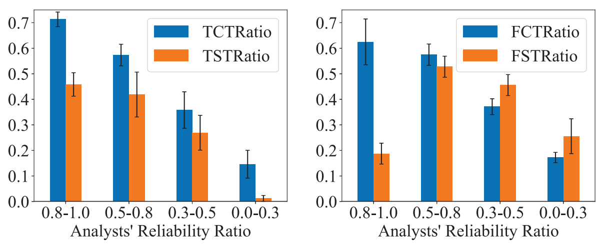
For the first question, Figure 4(a) plots the distributions of opinion and reliability of analysts' comments. We can see that approximately 65% comments are bullish, which means that analysts tend to express bullish opinions. Such a "buy side" opinion pattern is consistent with previous literature [1]. This phenomenon can be explained that analysts tend to motivate traders to buy stocks.

On the other hand, only half of comments are reliable, no matter which opinion polarities are expressed. It indicates that to reliably predict stock trends is not an easy task. Especially in Figure 4(b), only a tiny proportion of analysts have a reliability ratio larger than 0.8. In contrast, most of analysts' prediction accuracies are around 0.5. Here the reliability ratio refers to the proportion of reliable *comment units*, in which for each *comment unit* c_i, $r^{(c_i)} = 1$.



(a) Opinion Shift Ratio (b) *TSRatio* and *FSRatio*

Figure 5: Patterns of opinion shift of different analysts.



(a) Last prediction was correct (b) Last prediction was wrong

Figure 6: Reliability ratio of opinion shift behaviors.

3.2 Characteristics of Opinion Coherence

For the second question, Figure 5(a) gives the answer, in which analysts are divided into four groups based on their reliability ratios: *most reliable, reliable, unreliable* and *most unreliable.* It shows that most of analysts' average opinion shift ratios (*OSRatios*) are below 0.3. Here the *OSRatio* refers to the ratio of *comment units* posted by an analyst, in which each *comment unit* c_i expresses the opposite opinion polarity compared with its last *comment unit* c_{i-1}. That is, c_i and c_{i-1} belong to the same *comment sequence* and $o^{(c_i)} \neq o^{(c_{i-1})}$. Therefore, when discussing a stock, analysts tend to hold coherent opinions, rather than to shift opinions frequently. Besides, we observe that the analysts with higher reliability ratios have lower *OSRatios*, which means *OSRatio* may be considered as an indicator or feature when determining whether an analyst is reliable or not.

3.3 Patterns of Opinion Shift

For the last question, we analyze the relations between analysts' opinion shift behaviors and their reliability ratios.

3.3.1 When to Shift? We investigate under what conditions an analyst's opinion shift behavior occurs: after a successful prediction or a failed prediction. In other words, when the last *comment unit* c_{i-1} was separately reliable ($r^{(c_{i-1})} = 1$) and unreliable ($r^{(c_{i-1})} = 0$), we measure the probability that the current *comment unit* c_i expressed the opposite opinion polarity with c_{i-1} ($o^{(c_i)} \neq o^{(c_{i-1})}$). Figure 5(b) shows the observation results, in which the *TSRatio* (the ratio of True-then-Shift) and the *FSRatio* (the ratio of False-then-Shift) separately denotes the ratio of opinion shift behaviors when the last prediction was correct or incorrect. Generally, the *FSRatios* of analysts are larger than their *TSRatios.* It means that analysts are more likely to change their opinions about market trends if they made a wrong decision previously, especially for those analysts whose reliability ratios are below 0.3. However, the analysts whose reliability ratios are above 0.8 are exceptions: their *TSRatios* slightly exceed *FSRatios.* This phenomenon may be explained that instead of changing their views according to the previous predictions, they preferably insist on their own judgment about market trends. Therefore, *TSRatio* and *FSRatio* can be regarded as two indicators to distinguish reliable analysts from unreliable ones.

3.3.2 Is Shift a Good Strategy? We investigate if an opinion shift behavior will lead to a correct prediction. In Figure 6, there are two cases: the last prediction is correct or incorrect.

Firstly, when the last prediction was correct, Figure 6(a) shows that the *TCTRatio* is always larger than the *TSTRatio*, in which the *TCTRatio* (the reliability ratio of True-then-Constant) and the *TSTRatio* (the reliability ratio of True-then-Shift) separately denotes the reliability ratio of opinion hold or shift when the last prediction was correct. That is, in this case the opinion shift behavior is more likely to lead to an incorrect prediction. Thus, it is a good strategy for the analysts to keep the same opinion.

Secondly, when the last prediction was incorrect, Figure 6(b) shows that the analysts with different reliability ratios have different patterns: for the most reliable analysts, their *FCTRatios* (the reliability ratio of False-then-Constant) are much larger than *FSTRatios* (the reliability ratio of False-then-Shift). Besides, for those whose reliability ratios are between 0.5 and 0.8, their *FCTRatios* are slightly larger than *FSTRatios.* However, for those unreliable analysts whose reliability ratios are below 0.5, their *FCTRatios* are smaller than *FSTRatios.* Therefore, *FCTRatio* and *FSTRatio* can also be treated as two indicators to distinguish reliable analysts from unreliable ones.

The observations mentioned above demonstrate that there are evident relations between analysts' opinion dynamics and their reliability ratios. Therefore, for accurately modeling analyst's opinion dynamics, *OSRatio*, *TSRatio*, *FSRatio*, *FCTRatio* and *FSTRatio* mentioned above are considered as key features for predicting the reliability labels of stock comments in Table 2.

4 METHOD

Based on the observations above, this section presents the key components of our approach for modeling the reliability of stock comments. Figure 2 outlines the framework of the approach. Specifically, we first leverage FM model to classify the opinions polarity of historical *comment units* into bullish or bearish. Then, we construct *comment sequences* based on extracted comment opinions and time-series of historical stock prices. Moreover, we use a TSA model, ARMA, to predict future trends of stock prices. Next, we use an ensemble learning method, which incorporate SVM and ARMA together to predict the reliability label of each newly coming comment based on features extracted from *comment sequences* and price trends. Finally, we rank the comments based on their reliability value, which are used in the subsequent portfolio recommendation in Subsection 5.4.

4.1 Classification of Comment Opinions

The first step of the approach is the classification of comment opinions. The objective is to determine the opinion polarity $o^{(c_i)}$ of a *comment unit* c_i by analyzing the comment text $d^{(c_i)}$. The opinion polarity of each *comment unit* c_i will be used as a key feature in Subsection 4.3 to determine whether c_i is reliable or not.

Suppose that there are totally N *comment units* included in the dataset $\{(x_i, y_i)|i = 1, \cdots, N\}$, where $x_i \in \mathbb{R}^Q$ denote the extracted $tf - idf$ text features of the i-th *comment unit*; Q denotes the size of vocabulary, and $y_i \in \{-1, 1\}$ is the classification label, such that $y_i = -1$ means that the i-th *comment unit* is bearish and $y_i = 1$ means that the i-th *comment unit* is bullish. The prediction function of the classification model can be denoted as:

$$g : x \mapsto y = g(x). \tag{1}$$

Compared with some general text classification algorithms such as SVMs, we choose a 2-way Factorization Machine (FM with degree $d = 2$) [26] to implement the prediction function $g(\cdot)$. The reason is that FM not only alleviates the impact of high dimensionality of text features, but also captures some aspects of basic linguistics by modeling pairwise interactions of high-dimensional features with low-rank factorization [6].

Based on extracted features, the FM model is defined as:

$$\hat{g}(x) = w_0 + \sum_{i=1}^{Q} w_i x_i + \sum_{i=1}^{Q} \sum_{j=i+1}^{Q} \langle v_i, v_j \rangle x_i x_j. \tag{2}$$

The model parameters $\mathbf{\Theta} = \{\mathbf{w}, \mathbf{V}\}$: $\mathbf{w} \in \mathbb{R}^{Q+1}$, $\mathbf{V} \in \mathbb{R}^{Q \times k}$, k is the dimensionality of the factorization, and $\langle \cdot, \cdot \rangle$ denotes the dot product of two k-dimension vectors.

The left part of FM in Eq. (2) contains a bias and the unary interactions of each feature x_i with the target, which is similar to Logistic Regression (LR). However, the right part with two nested sums consists of all pairwise interactions of features x_i and x_j. Instead of using an independent parameter $\omega_{i,j}$, FM models the pairwise interactions between features using factorized parameters $\langle v_i, v_j \rangle$. Therefore, FM can estimate interactions even in problems with huge sparsity [26].

Back to the opinion classification task, modeling pairwise interactions and factorization parameterization are two main advantages of FM over the other classification models. The superiority is demonstrated in Subsection 5.2.

We use the stochastic gradient descent method with adaptive regularization [27] to learn the parameters of FM model. The hyperparameter k is tuned by cross validation. At last, the classification model is represented as:

$$g(x) = \begin{cases} 1 & \hat{g}(x) \geq 0 \\ -1 & \hat{g}(x) < 0 \end{cases}. \tag{3}$$

After the opinion polarity $o^{(c_i)}$ of each c_i has been determined by $o^{(c_i)} = g(d^{(c_i)})$, we calculate the reliability label $r^{(c_i)}$ for *comment sequences* construction. First, $t^{(c_i)}$ is used for aligning stock price data with stock comment data. Then, $r^{(c_i)}$ is computed by:

$$r^{(c_i)} = \begin{cases} 1 & o^{(c_i)} \cdot [sp(t^{(c_i)} + 1) - sp(t^{(c_i)})] \geq 0 \\ -1 & o^{(c_i)} \cdot [sp(t^{(c_i)} + 1) - sp(t^{(c_i)})] < 0 \end{cases}, \tag{4}$$

where $sp(t)$ denotes the stock price in the day t [5]. Eventually, *comment sequences* are constructed.

4.2 Prediction of Stock Time-Series

In addition to opinion polarities, the future trends of stock prices are also key factors to determine whether a *comment unit* is reliable. Furthermore, the stock prices as a kind of time-series data can be predicted to some extent by using Time-Series Analysis (TSA) models [3, 42]. Therefore, the second step of the approach is the time-series prediction of stock prices.

First, we use the stock prices to construct the time-series. Specifically, for each stock, the daily prices are used to construct one time-series, in which each node denotes the adjusted closing price of one day. Then, We use AutoRegressive Moving Average (ARMA), a well-known TSA model, to perform the prediction. Given a time-series of data X_t , the ARMA model is described as:

$$X_t = b + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}, \tag{5}$$

where $\varphi_1, \ldots, \varphi_p$ and $\theta_1, \ldots, \theta_q$ are the parameters; b is a constant; and ε_t is the i.i.d. random variable which is sampled from a normal distribution with zero mean, i.e. $\varepsilon_t \sim N(0, \sigma^2)$ where σ^2 is the variance of the distribution. In addition, p is the order of the autoregressive part and q is the order of the moving average part.

The parameters above can be computed by maximum likelihood estimation. Besides, p and q are tuned by using bayesian information criterion (BIC). After the ARMA model is trained on historical price series[6], for each *comment unit* c_i with the timestamp $t^{(c_i)}$, the model forecasts the stock price in the next trading day $\hat{sp}(t^{(c_i)} + 1)$ based on historical series $sp(t_1), sp(t_2), \cdots, sp(t^{(c_i)})$. Eventually, $\hat{sp}(c_i) = \hat{sp}(t^{(c_i)} + 1)$ and its standard error $err(c_i) = err(t^{(c_i)} + 1)$, i.e. variances of ε in Eq. (5), are treated as the features of stock trends in Table 2.

4.3 Ranking of Comment Reliability

The third step of the approach is the reliability ranking of stock comments. Firstly, we classify a *comment unit* c_i into reliable ($r^{(c_i)} = 1$) or unreliable ($r^{(c_i)} = -1$). After we obtain the opinion polarities of comments and trends of stock prices, and based on the observations in Section 3, we extract the corresponding multi-facet features and group them into four categories as shown in Table 2.

Based on the extracted features, we employ an ensemble learning method for reliability classification. Specifically, we combine SVMs and ARMA in a weighted voting classifier framework [13].

Note that here we choose SVMs, not FM in Subsection 4.1, because our other experiments (results omitted in this paper) show that the two models have comparable performances in reliability classification. This can be interpreted that the superiority of FM

[5]When the day t is not a trading day, it will be postponed to a nearest trading day.
[6]Note that ARMA model requires the input series data to be stationary, so we make a difference operation on the price series before feeding them into the ARMA model.

Table 2: List of extracted features from a *comment unit* $c_i = \{d^{(c_i)}, a^{(c_i)}, s^{(c_i)}, t^{(c_i)}, o^{(c_i)}, r^{(c_i)}\}$.

Category	Feature Descriptions
Opinion Polarity	The bullish or bearish polarity of the current comment
Historical Stock	# (bullish/bearish) comments about $s^{(c_i)}$ on $t^{(c_i)}$
Status*	# (bullish/bearish/reliable/unreliable) comments about $s^{(c_i)}$ over the past 7 days
Price Time	Prices sequence over the past 25 trading days
Series	The predicted price and its standard error in the next trading day by ARMA
Analysts'	# (bullish/bearish/reliable/unreliable) comments posted by $a^{(c_i)}$ over the past 7/30/90 days
Historical	# (bullish/bearish/reliable/unreliable) comments about $s^{(c_i)}$ by $a^{(c_i)}$ over the past 7/30/90 days
Performance	*OSRatio*, *TSRatio*, *FSRatio*, *FCTRatio* and *FSTRatio* of $a^{(c_i)}$ over the past 7/30/90 days

* This kind of features outline macro impact of public opinions on stock trends [8], aka. "word of mouth" effect.

is more prominent when the features are high-dimensional and sparse, such as the text features in Subsection 4.1.

Specifically, in terms of the SVM part, the radial basis function (RBF) kernel, $\kappa(x_1, x_2)) = \langle\phi(x_1), \phi(x_2)\rangle = e^{-\gamma|x_1-x_2|^2}$, is used. Note that the function $\phi(\cdot)$ maps the original features into a high dimensional kernel space where the optimal decision hyperplane $\hat{h_1}(c_i) = \langle\omega, \phi(c_i)\rangle + b$ can be computed.

To compute the optimal ω and b in $\hat{h_1}(\cdot)$, we optimize:

$$\begin{aligned} &\min_{\omega,b,\xi} \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{N}\xi_i, \\ &s.t. \quad y_i(\omega^T\phi(c_i) + b) \geq 1 - \xi_i, \\ &\qquad \xi_i \geq 0, i = 1, \cdots, N, \end{aligned} \tag{6}$$

where C trades off the misclassification of training examples against simplicity of the decision hyperplane.

For the ARMA part, we use the prediction results in Subsection 4.2 to calculate the classification function:

$$\hat{h_2}(c_i) = \begin{cases} \frac{1}{err(c_i)} & \left[\hat{sp}(t^{(c_i)}+1) - sp(t^{(c_i)})\right] \cdot o^{(c_i)} \geq 0 \\ -\frac{1}{err(c_i)} & \left[\hat{sp}(t^{(c_i)}+1) - sp(t^{(c_i)})\right] \cdot o^{(c_i)} < 0 \end{cases}, \tag{7}$$

where $1/err(c_i)$ is treated as the classification probability of the ARMA model, and $t^{(c_i)}$ and $o^{(c_i)}$ separately denotes the date and opinion polarity of the *comment unit x*.

Therefore, the final classification model is represented as:

$$h(c_i) = \begin{cases} 1 & \hat{h}(c_i) \geq 0 \\ -1 & \hat{h}(c_i) < 0 \end{cases}, \tag{8}$$

$$\hat{h}(c_i) = u\hat{h_1}(c_i) + (1-u)\hat{h_2}(c_i), \tag{9}$$

where $u \in [0, 1]$ is a scaling factor to determine the weights of $\hat{h_1}(c_i)$ and $\hat{h_2}(c_i)$. Despite of its simplicity, we find this linear combination always obtains good results in practice. We test the performance of our method using different u in Section 5.6. Besides, we have tried to use other nonlinear combination options, including logarithm, quadratic or normalization operation on $\hat{h_1}(c_i)$ and $\hat{h_2}(c_i)$ before they are combined together. None of these alternatives could improve performance dramatically.

Please note that C in Eq. (6), u in Eq. (9) and γ in the kernel function of SVM are tuned by preliminary experiments.

Lastly, the *comment units* with the predicted reliability labels are ranked by measuring the classification probability: $rv(c_i) = |\hat{h}(c_i)|$; we call it *reliability value* or *confidence*. Note that $\hat{h}(\cdot)$ denotes the classification function in Eq. (9). That is, a comment c_i with higher *reliability value*, i.e. $rv(c_i)$, is more likely to be a reliable comment if $r^{(c_i)} = 1$. On the other hand, if $r^{(c_i)} = -1$, c_i with higher $rv(c_i)$ is more likely to be a unreliable comment.

5 EXPERIMENTS

In this section, we conduct extensive experiments on real-world data, and provide a systematic trading simulation to evaluate the effectiveness of our method.

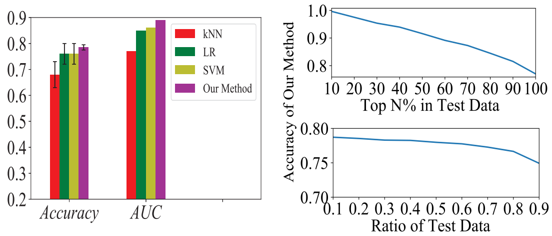
5.1 Experiment Settings

Evaluation Tasks. We conduct three tasks: the evaluation on opinion classification, the evaluation on reliability prediction, and a market simulation. These tasks are separately denoted as Task 1, 2 and 3. Besides, we present two additional experiments to further demonstrate the superiority of our multi-facet features selection and ensemble learning based framework.

Datasets. The datasets [7] have been described in Section 2. Ten participants having professional stock knowledge are employed to label the opinion polarities of each comment. For Tasks 1 and 2, we split the datasets into the training set (90%) and the test set (10%). In Task 1, the test set is randomly sampled and 10-fold cross validation is performed. In Task 2, the 10% with the latest timestamp is the test set. In Task 3, the training set covers 16 months (from Aug 2014 to Dec 2015), and the test set comes from the following 12 months (from Dec 2015 to Dec 2016). Figure 10 shows that there are two market stages in the test set: the bearish market before May and the bullish market after June.

Baselines. The baselines include three state-of-the-art stock prediction models: MKL-SVM [20], TSLDA [22] and SFM [42]; three classification models: SVM, Logistic Regression (LR) and k-Nearest Neighbor (kNN); and two TSA models: ARMA and Generalized AutoRegressive Conditional Heteroskedasticity (GARCH). Note that kNN is evaluated only in Task 1; MKL-SVM, TSLDA, SFM, ARMA and GARCH are evaluated only in Tasks 2 and 3.

[7] https://github.com/zhangchen010295/Reliability-Modeling-for-Stock-Comments



(a) Accuracy Evaluation (b) Robustness Evaluation

Figure 7: Evaluation on opinion classification.

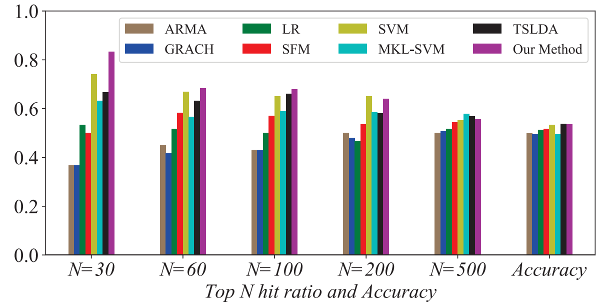


Figure 8: Evaluation on reliability prediction.

Parameters tuning. For all the methods, we tune their parameters using grid search with cross-validation by further dividing the training set into 90% for model fitting and 10% for parameter optimization. Moreover, the parameters of two TSA models are tuned by maximum likelihood estimation. Specifically, we use Augmented Dickey-Fuller test (ADF)[8] to ensure that the time-series used are stationary. Then, bayesian information criterion (BIC) is adopted, when choosing the proper order, to guarantee the sufficiency of two TSA models. Lastly, we choose the optimal results from each model to compare.

Evaluation metrics. In Task 1, we use *Accuracy* (the ratio of correctly identified bullish and bearish comments) and *AUC* (Area Under the Receiver Operating Characteristic Curve with macro average) to evaluate all the methods. In Task 2, *Accuracy* (the ratio of correctly identified reliable and unreliable comments divided) and *Top N hit ratio* [9] are chosen as the criteria. In Task 3, *Profit* and *Sharpe ratio* are used as the performance indicators. Note that *Sharpe ratio* [10] is a more financially intuitive metric to examine the performance of an investment by adjusting for its risk [20].

5.2 Evaluation on Opinion Classification

For opinion classification task, the same $tf - idf$ text features extracted from the stock comments, and L2 norm regularization based feature selection, are used in each model. Figure 7(a) shows that our method achieves the best performance and the lowest standard error. Such superiority is due to the usage of the FM model when extracting opinion polarities from the high-dimensional text features of stock comments.

Figure 7(b) demonstrates the robustness evaluation of our method, in which the figure on the top shows that the performance of our method can be further improved if we sort the samples according to the classification probability and choose top N% samples for validation; the figure on the bottom shows the stability of the performance of our method when we increase the ratio of test data.

5.3 Evaluation on Reliability Prediction

For reliability prediction task, MKL-SVM and TSLDA use the features of comment text and stock time-series. SVM and LR use the features in Table 2. Besides, the feature selection with L2 norm regularization is used for obtaining the optimal results. For SFM, ARMA and GARCH, the features of stock time-series are used. In addition, their performance of up/down trend prediction on stock prices are evaluated here.

Figure 8 lists the results. Although the *accuracy* of our method is comparable with the baselines, which are both around 0.5, the *top N hit ratio* of our method is better than that of the baselines, in which N is set to 30, 60, 100, 200 and 500. In other words, our method has superior ranking ability. As will be demonstrated in Subsections 5.5 and 5.6, such improvement comes from two factors: the multi-facet features selection and the ensemble learning framework. As a consequence, our method can discover the most reliable comments more effectively.

5.4 Market Trading Simulation

To demonstrate the application of our method, we simulate a stock trading in 12-month period on a virtual trading platform[5]. The daily capital is 10,000 CNY (Chinese yuan).

5.4.1 Stock Trading Strategies. Based on the ranked stock comments in Subsection 4.3, we design two types of strategies. First, for each trading day we build a stock pool $P = s_1, s_2, \cdots, s_{|P|}$, in which each stock is supposed to go up in the next trading day. The up/down trend of a stock s_j is calculated by:

$$trend(s_j) = \begin{cases} up & cf(s_j) \geq 0 \\ down & cf(s_j) < 0 \end{cases}, \tag{10}$$

$$cf(s_j) = \sum_{c_i \in C_{s_j}} \frac{o^{(c_i)} \cdot r^{(c_i)} \cdot rv(c_i)}{|C_{s_j}|}, \tag{11}$$

where C_{s_j} denotes a set of stock comments on s_j. Eq. (10) means that if most of the reliable comments reveal bullish opinions on s_j and/or most unreliable comments reveal bearish opinions, s_j is supposed to increase. Besides, each stock in P is assigned a weight: $w(s_j) = |cf(s_j)|$.

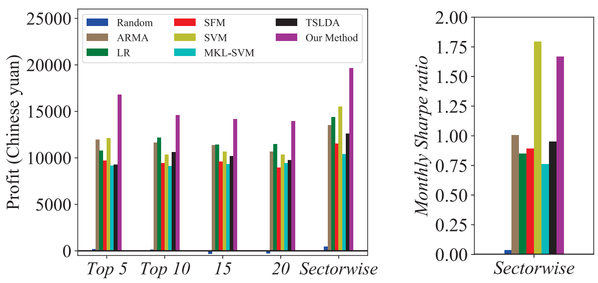
Second, we pick K stocks from P to invest. We invest in 10,000 CNY on these stocks at the opening price; and allocate uniform money to each stock. After the purchase, we hold the stocks for one day and sell them at the closing price. We also assume a zero transaction cost which is consistent with the prior studies [6, 18, 21].

Last, we explain two strategies of how we pick K stocks.

[8]https://en.wikipedia.org/wiki/Augmented_Dickey-Fuller_test.

[9]We get *Top N hit ratio* by ranking the test samples based on their *reliability value*, and then checking the *Accuracy* on top N samples. For ARMA and GARCH, the *reliability value* is measured as the inverse of predicted standard error $err(c_i)$ in Eq. (7).

[10]https://en.wikipedia.org /wiki/Sharpe_ratio.



(a) Profits (b) *Sharpe ratio*

Figure 9: Results of the market simulation.

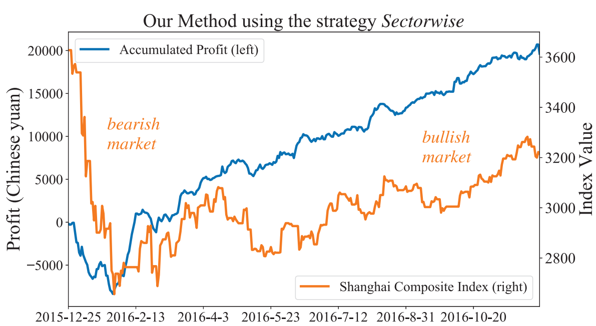


Figure 10: Accumulated profit of our method.

(1) *Top K*: We sort stocks in P based on their weights and pick top K stocks.
(2) *Sectorwise*: First, the money is uniformly divided into Se fold, where Se denotes the number of sectors (see Table 1). Second, for each sector we pick a stock s_j with the highest weight in this sector and allocate each fold of money to s_j.

5.4.2 Simulation Results. Figure 9 shows the simulation results [11]: our method performs significantly better than the baselines. Note that to verify the statistical significance of our results, we also perform a random test by randomly buying or selling for 1,000 trials. As shown in Figure 9(a), the mean profit over the random test is close to zero. Besides, the superiority of the diversified portfolio strategy is demonstrated that the profit of using *Sectorwise* is consistently better than that of using *Top K*, in which K is set to 5, 10, 15 and 20, respectively. Moreover, Figure 9(b) indicates that our method achieves a higher monthly *Sharpe ratio* than all of the baselines except SVM [12], which means that our method can provide better return for the same risk.

Especially, the accumulated profit in Figure 10 shows that our method is robust under different market environments, such as the bearish market before May and the bullish market after June. For example, in an abrupt bearish market before February, the loss of our method is much lower than *Shanghai Composite Index*. Note that *Shanghai Composite Index* in Figure 10 indicates the whole trend of Chinese stock market.

[11]We omit the result of GARCH because the performance of GARCH is quite close to ARMA, such as *Profit* and *Sharpe ratio*.
[12]Subsection 5.6 gives the explanation.

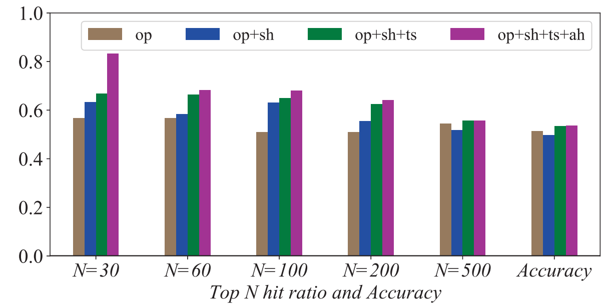


Figure 11: Performance of our method with different feature combinations on reliability prediction. Note that *op*, *sh*, *ts* and *ah* separately denotes the features of opinion polarity, historical stock status, time-series of stock price and analysts' historical performance mentioned in Table 2.

Table 3: Performance of our method with feature ablations.

	Top N hit ratio and Accuracy					
	N=30	60	100	200	500	*Accuracy*
Our method	**0.833**	**0.683**	**0.68**	**0.64**	**0.556**	**0.536**
$-ts$	0.567	0.483	0.54	0.57	0.524	0.509
$-ah$	0.667	0.65	0.65	0.625	0.556	0.535
$-op$	0.7	0.65	0.62	0.595	0.51	0.507
$-sh$	0.8	0.617	0.56	0.575	0.57	0.524

5.5 Superiority of Multi-facet Feature Selection

To analyze the discriminative power of the used features in Table 2, we evaluate the performance of our method with different feature combinations. As shown in Figure 11, the accuracy of the feature *op* [13] on reliability prediction is improved by the features *sh*, *ts* and *ah*. Particularly, the contribution rate of different features on such improvement can be ranked as $ah > ts > sh$.

Besides, in order to further understand the features that are responsible for the performance of our method, we perform a feature ablation where we remove one group of features from our model at a time. The results, shown in Table 3, indicate that *ts* and *ah* features are most important.

5.6 Advantage of Ensemble-based Framework

Figure 9(b) indicates that our method achieves a higher monthly *Sharpe ratio* than all of the baselines except SVM. It is interesting to find that although SVM and our method have comparable monthly *Sharpe ratio*, our method performs better than SVM in the metric of *profit* as shown in Figure9(a). This is because although ensembling ARMA into SVM slightly decreases the *Sharpe ratio*, ARMA has the advantage in predicting the higher daily volatility in the stock

[13]When *op* is used alone, we actually use the $tf-idf$ text features of comments because the opinion polarity is an unidimensional feature.

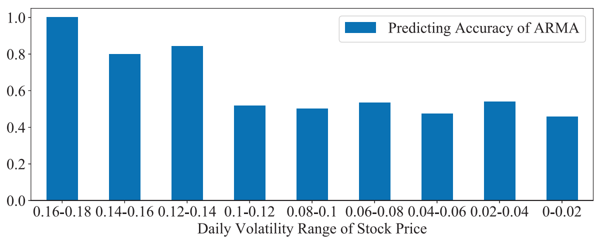


Figure 12: Accuracy of ARMA on various volatilities.

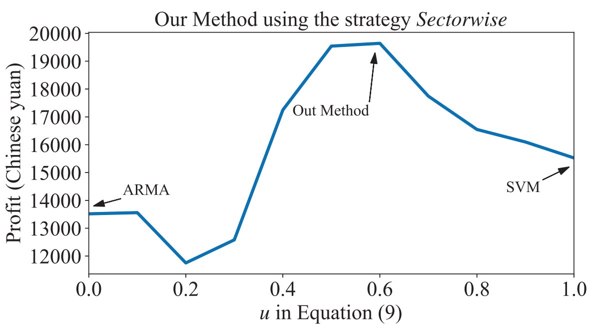


Figure 13: Profit of our method with different u.

market while it makes mistakes for the lower daily volatility, as shown in Figure 12.

In addition, to further validate the superiority of our ensemble learning framework, we perform simulations on the framework equipped with various u in Eq. (9), which controls the weights between SVM and ARMA[14]. The simulation results[15] are shown in Figure 13, where the strategy *Sectorwise* is adopted. It is clear that combining SVM and ARMA in a unified framework can bring more profits. Specifically, when $u > 0.4$, the performance of the framework exceeds both SVM and ARMA. In other words, our method has superior ranking performance due to the used ensemble learning framework. As a consequence, our method can be utilized to build profitable trading strategies.

6 RELATED WORK

Opinion Mining of Stock Comments. Recently, many studies have been trying to extract opinion polarities of investors and analyze the impact the opinions have on market movements [7, 10]. Corredor et al. [8] used stock characteristics, such as stock turnover and trading volume, to discover investors' opinions in four European stock markets. The results indicated that opinion polarities held by investors have a significant influence on stock returns. More researches employed Sentiment Analysis (SA) to extract investors' opinions. For example, Porshnev et al. [25] extracted investors' opinion polarities from stock comments on Twitter by SVM and predicted DJIA and S&P500 indicators by neural network. Antweiler and Frank [1] performed sentiment analysis to discover bullishness index from stock comments posted on Yahoo! Finance.

In short, previous studies basically focused on identifying coarse-grained opinions. However, few studies made a fine-grained analysis for distinguishing reliable comments. Therefore, we propose to model the reliability of comments.

Stock Price Prediction. Stock price prediction has attracted much attention in the fields of finance and computer science. In general, historical stock time-series can be utilized to predict future movements [3]. Some TSA models, such as ARMA, GARCH and their variants, were developed to model and predict stock prices [30]. Zhang et al. [42] proposed a State Frequency Memory (SFM) recurrent network model to predict stock prices by capturing the long and short term trading patterns from historical stock time-series. This regression model can learn and fuse the multi-frequency patterns underlying the price time-series so as to make short and long term price predictions. Besides, some studies have incorporated financial news to improve prediction by using natural language processing and machine learning models [5]. In general, text features such as bags-of-words, noun phrases, and named entities were extracted from news documents, and classification or regression models were adopted. Wang and Hua [35] proposed a semiparametric Gaussian copula regression model to predict financial risks. Luss et al. [20] used SVM with multiple kernel learning to predict abnormal returns from news and equity returns. Ding et al. [15] exploited financial news to extract structured event features to predict stock trends. Compared with explicit semantic features in financial news, another kind of implicit features, i.e. sentiments or opinion polarities, are also significant for stock market prediction [4, 12, 29, 37]. Bollen et al. [4] detected large-scale collective emotions representing public opinions on Twitter and found that the public opinions are correlated with the volatility of DJIA. Nguyen et al. [22] proposed a new topic model TSLDA to capture topics and their sentiments from social media for stock prediction. Si et al. [29] proposed a model to regress topic-sentiment time-series and stock's price time-series. There are some other studies that combined time-series data and text data, such as social media [16, 36, 43], news [18] and stock message board [19, 23, 28], to enhance prediction performance.

Therefore, although previous studies used opinion to predict stock movement, few studies considered the hidden opinion dynamics. To this end, we propose an approach to predict the reliability of stock comments by exploiting multi-facet features.

7 CONCLUSION

We developed an approach to model the reliability of stock comments. Along this line, we first analyzed stock comments in a time-evolving manner, and uncovered some important findings regarding the coherent characteristics of analysts' opinions and the opinion shifting patterns. Then, we extracted multi-facet features from stock prices, comment text, and analyst behaviors. These features are critical for modeling the reliability of stock comments. Next, we designed a hybrid strategy to measure the reliability of each comment by combining the power of SVM and ARMA in an ensemble learning based framework. Finally, we demonstrated the effectiveness of our model for identifying reliable and unreliable comments through

[14] In the previous experiments, u in our method is tuned as 0.59.
[15] Note that when $u = 1$, the framework is degenerated as SVM; when $u = 0$, the framework is degenerated as ARMA.

extensive experiments on real-world data. In particular, we showed that the identified comments could be exploited for building trading strategies and produced tangible profits in our trading simulation.

Further work can focus on the extension of the approach by using Recurrent Neural Networks (RNN) to exploit temporal dependency of successive comments. Besides, how to predict weekly and monthly stock trends based on comments analysis remain to be a challenge, because currently we only consider daily prediction and assume opinions in comments are short-term. Also, how to extend the approach to model analyst's profile based on their comments, such as measuring analyst's similarity for semantic clustering, is an interesting but challenging problem. Lastly, our approach will be applicable for various applications, such as intelligent recommender [31, 41], financial event detection [34, 38, 39], opinion mining [45], business opportunity discovery [32, 33], and user profiling [9, 40].

ACKNOWLEDGMENTS

This work is supported by National Basic Research Program of China (2013CB329305) and Natural Science Foundation of China (61402447, 61502466, 61672501). This work is also sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

REFERENCES

[1] Werner Antweiler and Murray Z Frank. 2004. Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance* 59, 3 (2004), 1259–1294.
[2] Paul Asquith, Michael B Mikhail, and Andrea S Au. 2005. Information content of equity analyst reports. *Journal of financial economics* 75, 2 (2005), 245–282.
[3] George S Atsalakis and Kimon P Valavanis. 2009. Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert Systems with Applications* 36, 7 (2009), 10696–10707.
[4] Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of computational science* 2, 1 (2011), 1–8.
[5] Sunandan Chakraborty, Ashwin Venkataraman, Srikanth Jagabathula, and Lakshminarayanan Subramanian. 2016. Predicting Socio-Economic Indicators using News Events. In *SIGKDD*. ACM, 1455–1464.
[6] Chen Chen, Wu Dongxing, Hou Chunyan, and Yuan Xiaojie. 2014. Exploiting social media for stock market prediction with factorization machine. In *WI*. IEEE Computer Society, 142–149.
[7] Hailiang Chen, Prabuddha De, Yu Hu, and Byoung-Hyoun Hwang. 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies* 27, 5 (2014), 1367–1403.
[8] Pilar Corredor, Elena Ferrer, and Rafael Santamaria. 2013. Investor sentiment effect in stock markets: Stock characteristics or country-specific factors? *International Review of Economics & Finance* 27 (2013), 572–591.
[9] Peng Cui, Fei Wang, Shaowei Liu, Mingdong Ou, Shiqiang Yang, and Lifeng Sun. 2011. Who should share what?: item-level social influence prediction for users and posts ranking. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. ACM, 185–194.
[10] Zhi Da, Joseph Engelberg, and Pengjie Gao. 2011. In search of attention. *The Journal of Finance* 66, 5 (2011), 1461–1499.
[11] Krishnan Dandapani, Sanjiv Sabherwal, Salil K Sarkar, and Ying Zhang. 2008. Online talk: does it matter? *Managerial Finance* 34, 6 (2008), 423–436.
[12] Sanjiv R Das and Mike Y Chen. 2007. Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management science* 53, 9 (2007), 1375–1388.
[13] Thomas G Dietterich. 2000. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*. Springer, 1–15.
[14] Xiao Ding, Yue Zhang, Ting Liu, and Junwen Duan. 2014. Using Structured Events to Predict Stock Price Movement: An Empirical Investigation.. In *EMNLP*. 1415–1425.
[15] Xiao Ding, Yue Zhang, Ting Liu, and Junwen Duan. 2015. Deep Learning for Event-Driven Stock Prediction. In *IJCAI*. 2327–2333.
[16] Eric Gilbert and Karrie Karahalios. 2010. Widespread Worry and the Stock Market.. In *ICWSM*. 59–65.
[17] György Komaromi. 2002. Why have the stock markets become noisier after the revolution of information and technology. In *Evolutions of Institutions and the Knowledge Economy Conference Proceedings. University of Debrecen.*
[18] Victor Lavrenko, Matt Schmill, Dawn Lawrie, Paul Ogilvie, David Jensen, and James Allan. 2000. Mining of concurrent text and time series. In *KDD-2000 Workshop on Text Mining*. 37–44.
[19] Banghui Luo, Jianping Zeng, and Jiangjiao Duan. 2016. Emotion space model for classifying opinions in stock message board. *Expert Systems with Applications* 44 (2016), 138–146.
[20] Ronny Luss and Alexandre d'Aspremont. 2015. Predicting abnormal returns from news using text classification. *Quantitative Finance* 15, 6 (2015), 999–1012.
[21] M-A Mittermayer. 2004. Forecasting intraday stock price trends with text mining techniques. In *HICSS*. IEEE, 10–pp.
[22] Thien Hai Nguyen and Kiyoaki Shirai. 2015. Topic Modeling based Sentiment Analysis on Social Media for Stock Market Prediction.. In *ACL (1)*. 1354–1364.
[23] Thien Hai Nguyen, Kiyoaki Shirai, and Julien Velcin. 2015. Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications* 42, 24 (2015), 9603–9611.
[24] Natalie Y Oh, Jerry T Parwada, and Terry S Walter. 2004. Online investorsâĂŹ trading behaviour and performance: Evidence from the Korean equity market. *University of New South Wales, Australia* (2004).
[25] Alexander Porshnev, Ilya Redkin, and Alexey Shevchenko. 2013. Machine learning in prediction of stock market indicators based on historical data and data from twitter sentiment analysis. In *ICDM Workshops*. IEEE, 440–444.
[26] Steffen Rendle. 2010. Factorization machines. In *ICDM*. IEEE, 995–1000.
[27] Steffen Rendle. 2012. Learning recommender systems with adaptive regularization. In *WSDM*. ACM, 133–142.
[28] Vivek Sehgal and Charles Song. 2007. Sops: stock prediction using web sentiment. In *ICDM Workshops*. IEEE, 21–26.
[29] Jianfeng Si, Arjun Mukherjee, Bing Liu, Sinno Jialin Pan, Qing Li, and Huayi Li. 2014. Exploiting Social Relations and Sentiment for Stock Prediction.. In *EMNLP*, Vol. 14. 1139–1145.
[30] Stephen J Taylor. 2008. *Modelling financial time series*. world scientific.
[31] Hao Wang, Yanmei Fu, Qinyong Wang, Hongzhi Yin, Changying Du, and Hui Xiong. 2017. A location-sentiment-aware recommender system for both hometown and out-of-town users. In *KDD*. ACM, 1135–1143.
[32] Hao Wang and Yukio Ohsawa. 2013. Idea discovery: A scenario-based systematic approach for decision making in market innovation. *Expert Systems with Applications* 40, 2 (2013), 429–438.
[33] Hao Wang, Yukio Ohsawa, and Yoko Nishihara. 2012. Innovation support system for creative product design based on chance discovery. *Expert Systems with Applications* 39, 5 (2012), 4890–4897.
[34] Wei Wang, Chen Zhang, Hao Wang, Yang Gao, and Yuanman Zheng. 2017. A cognition graph approach for insights generation from event sequences. *Cluster Computing* 20, 2 (2017), 1679–1690.
[35] William Yang Wang and Zhenhao Hua. 2014. A Semiparametric Gaussian Copula Regression Model for Predicting Financial Risks from Earnings Calls.. In *ACL (1)*. 1155–1165.
[36] Steve Y Yang, Sheung Yin Kevin Mo, and Anqi Liu. 2015. Twitter financial community sentiment and its predictive relationship to stock market movement. *Quantitative Finance* 15, 10 (2015), 1637–1656.
[37] Akira Yoshihara, Kazuhiro Seki, and Kuniaki Uehara. 2015. Leveraging temporal properties of news events for stock market prediction. *Artificial Intelligence Research* 5, 1 (2015), 103.
[38] Chen Zhang, Hao Wang, Liangliang Cao, Wei Wang, and Fanjiang Xu. 2016. A hybrid term–term relations analysis approach for topic detection. *Knowledge-Based Systems* 93 (2016), 109–120.
[39] Chen Zhang, Hao Wang, Wei Wang, and Fanjiang Xu. 2015. Rcfged: retrospective coarse and fine-grained event detection from online news. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*. IEEE, 139–144.
[40] Chen Zhang, Hao Wang, and Yingcai Wu. 2018. ResumeVis: A Visual Analytics System to Discover Semantic Information in Semi-structured Resume Data. *ACM TIST* (2018).
[41] Chen Zhang, Hao Wang, and Hui Xiong. 2017. An Automatic Approach for Transit Advertising in Public Transportation Systems. In *ICDM*. IEEE, 1183–1188.
[42] Liheng Zhang, Charu Aggarwal, and Guo-Jun Qi. 2017. Stock Price Prediction via Discovering Multi-Frequency Trading Patterns. In *SIGKDD*. ACM, 2141–2149.
[43] Li Zhang, Keli Xiao, Qi Liu, Yefan Tao, and Yuefan Deng. 2015. Modeling Social Attention for Stock Analysis: An Influence Propagation Perspective. In *ICDM*. IEEE, 609–618.
[44] Ying Zhang and Peggy E Swanson. 2010. Are day traders bias free?-evidence from internet stock message boards. *Journal of Economics and Finance* 34, 1 (2010), 96–112.
[45] Qiyun Zhao, Hao Wang, Pin Lv, and Chen Zhang. 2014. A bootstrapping based refinement framework for mining opinion words and targets. In *CIKM*. ACM, 1995–1998.