Stock Price Forecasting with Support Vector Machines based on Web Financial Information Sentiment Analysis

Run Cao, Xun Liang, and Zhihao Ni

School of Information, Renmin University of China, Beijing, P. R. China
{caorun,xliang,alexni}@ruc.edu.cn

Abstract. The stock price forecasting has always been considered as a difficult problem in time series prediction. Mass of financial Internet information play an important role in the financial markets, information sentiment is an important indicator reflecting the ideas and emotions of investors and traders. Most of the existing research use the stock's historical price and technical indicators to predict future price trends of the stock without taking the impact of financial information into account. In this paper, we further explore the relationship between the Internet financial information and financial markets, including the relations between the Internet financial information content, Internet financial information sentimental value and stock price. We collect the news of three stocks in the Chinese stock market in the GEM in a few large portals, use the text sentiment analysis algorithm to calculate the sentimental value of the corresponding Internet financial information, combined with the stock price data, implant Support Vector Machines to analyzes and forecasts on the stock price, the accuracy of the prediction of stock prices has been improved.

Keywords. Text sentiment analysis, SVM, Stock price prediction.

1 Introduction

The financial market has always been a research hotspot, and stock price forecasting has been considered the most difficult challenges in the time series prediction. Stock price time series are data intensive, noisy, dynamic, unstructured, and have a high degree of uncertainty [1]. Stock price time series is more like a random walk curve [2], using the traditional statistical methods to predict stock price changes has been proven to be very difficult.

With the rapid development of communication technology and the Internet, people can access to the stock news and reviews through the Internet anytime, anywhere. Although the Internet is just one of the channels for an investor to access the financial information, and often overlap with newspapers, television and other media, but Internet information is more quickly and comprehensive, it could cover other channels, financial Internet information play an important role in the financial markets.

Most of financial information is unstructured text, the main forms are financial news on the major portals, financial analyst's comments, national policies and announcements, and the online community discussion. According to the efficient market hypothesis, the financial information has an important impact on the financial market volatility. The information has two dimensions characteristics: information volume and information sentiment, the information sentiment analysis has become an important topic of natural language processing and machine intelligence field.

In the financial markets, information sentiment is an important indicator reflecting the opinions and emotions of investors and traders. Financial information integrates a variety of factors in the market and affects the stock market participants' investment trading behavior to a large extent. The quantify effectiveness of the financial information can be confirmed directly through financial model, combining with existing financial mathematical model. Previous studies have shown that Internet financial information volume and stock market volatility are closely related [3], [4], [5], [6].

In this paper, we further explore the relationship between the Internet financial information volume, Internet financial information sentimental value and the stock price. We have three stocks in the Chinese stock market in the GEM as examples, the news of these three stocks in a few large portals have been collected， then we use the text sentiment analysis algorithm to calculate the sentimental value of the corresponding Internet financial information, combined with the stock price data, implant Support Vector Machines to analyzes and forecasts the stock price, the accuracy of the prediction of stock prices has been improved.

2 Related work

2.1 Stock Price Forecasting

The stock price forecasting can be divided into two branches, the industry and academia. Traditional industry is mainly based on fundamental analysis and technical analysis. Academics are mainly expanding the stock price forecasts from two aspects: statistical methods and machine learning method. Methods based on statistical are mainly using time series analysis methods, including linear regression prediction, polynomial regression prediction, ARMA modeling, GARCH modeling and so on. Methods based on machine learning often use non-linear prediction and intelligent learning, including the Grey Theory [7], Artificial Neural Networks [8], Support vector machine [9], Markov model, fuzzy network [11]and other.

Methods based on machine learning can solve the noise data problem, learning nonlinear models efficiently, as the stock market is a typical nonlinear complex system. In recent years, the neural network has been successfully applied to financial time series modeling from Stock Price Index [8], [12], to the option price [13]. Unlike the traditional statistical methods, neural networks are data-driven, does not require assumptions and parameters, nonlinear system can be well fitted [14], but the neural network optimization may fall into the Local optimum. The Support Vector Machine developed by Vapnik solve this problem successfully, because of the principle of structural risk optimal, SVM was able to reach the global optimum, and also overcoming the over fitting problem, which making SVM improved a lot compared to ANN in stock price prediction [15], [16] ,[17].

With the development of machine learning methods, integrated forecasting methods have got more and more attention. For example Md.Rafiul and Hassan applied an integration of HMM, ANN, GA model to predict the stock market [18]. Bildirici and

Erisn use ANN to restructure on GARCH model, and applied to the Turkish stock market in order to verify the validity of the hybrid model [19]. Wen use Box theory and SVM to build an automatic stock Decision Support System [20].

In machine learning, we focus on the input and output of the data collection. At present, for predicting stock prices, the input often includes a variety of macroeconomic indicators, and stock historical transaction data. In this article, the Internet Financial News as another exogenous input is introduced into the prediction of stock price and thus to improve the prediction accuracy.

2.2 Text sentiment analysis

Text sentiment analysis, also known as Opinion Mining, we can acquire the opinion of the text through the analyzed information sentiment tendency. For example, by automatically analyzing the text content of the comments of a commodity, we can get the consumer's attitudes and opinions of the goods. At present, the application of sentiment analysis mainly focused on: user comments analysis, decision-making, monitoring public opinion, and the information prediction. Text sentiment analysis is mainly from two aspects, sentimental knowledge-based approach and the method based on feature classification. The former use the existing sentiment dictionary or industry dictionary to calculate and obtain the polarity of the text; while the method based on feature classification use machine learning methods by selecting a large number of significant features to complete the classification task [21].

Present research are mainly use text sentiment analysis in merchandise online reviews, including books, movies, stocks, and electronic products. Mainly use empirical analysis, explore how the sentimental trend of online reviews will affect consumers' purchasing behavior and how to affect the mechanism of the related product sales, and establish the theoretical model [22], [23], [24]. Text sentiment analysis has also gradually attracted the attention of many Internet companies. Google has already applied text sentiment analysis into the search engine to provide users with more effective and versatile service. At the same time, the Internet video site YouTube also take text sentiment analysis into application in 2008, each user's comments are divided into "Good Comment" and "Poor Comment", allows users to glance on the previous attitude of the audience on the current video based on this statistics.

Although the Internet text analysis has become the common interest of many researchers, but the analysis specifically on the financial text are still rare. Devitt and other people use the of the Financial Review text sentimental polarity recognition, make predictions to future financial trends [25]. Das and Chen use linear regression method, find that the stock index and online stock analysis sentiment are significant positive correlation, but for a single company, this relationship is not established [6]. Antweiler and Frank discussed the relationship between the number of online stock analysis, stock analysis sentiment and stock trading volume, volatility and other indicators, the significant correlation between some variables show that the influence of online stock analysis on the stock market [3]. In this paper, we apply the text sentiment analysis in individual company news and stock analysts, daily information volume and financial information sentimental value are collected, then a machine learning method is used to predict the ups and downs of the stock price. This study is very innovative and has a strong practical significance.

3 Model and Method

3.1 The framework of the method

The framework of the overall project is depicted in Fig.1. with four components. The first component is the collection and computing of experimental information, including the financial text information and stock price information; The second component is the experimental input data preprocessing, including transfer the input data into Libsvm data input format, establish training set and test set, and normalized all input values; The third one is group experiments, all three stocks are divided into four groups separately, conducted SVM training, training the model to predict stock prices, and tested on the test set; The fourth component is calculate the evaluation score of the metrics to measure the performance of the experiment results data .

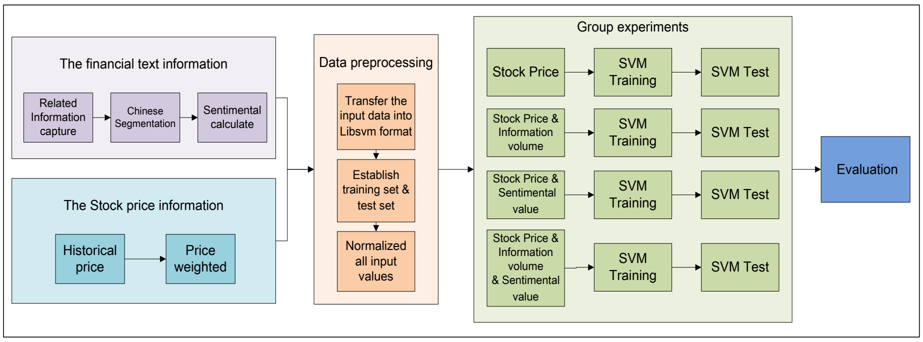

Fig. 1. The experiment framework

3.2 The calculation of Financial News sentimental value

For the Internet text, external factors and internal factors determine their nature. External factors are like the number of text on a particular topic appeared on the Internet within a specified time, and internal factors describe the nature of a single text. For a single text, its nature depends on its content and intensity, content refers to the theme, time, style of the text and intensity refers to the impact factor of the text.

The sentiment of financial text corresponds to the tendency of the attitude embodied in the text, i.e., bullish, bearish or neutral, and the financial text intensity reflects the influence of the text. High intensity of the text has a greater influence on financial markets, the low intensity of the text for the influence of the financial market is relatively weak. A sentimental value that is calculated according to a financial text, then the positive and negative symbols of the value means bullish bearish or neutral, while the absolute value represent the intensity of the text.

In this experiment, we use the algorithm Nan Li and Desheng Dash Wu proposed in 2010[26], the text sentiment algorithm based on Hownet sentiment dictionary. Assume that the current Financial News *p*, first conduct the segmentation tool to convert it into a sequence constituted by the word, that is, $\{w_1, w_2, w_3, \ldots w_n\}$, the number of total words is n, on each one the $w_i (i=1, 2, 3, \ldots n)$ to calculate an sentimental value of v_i, then the sentimental value of the entire Financial News *p* is the sum of all the words sentimental value. The Financial News sentimental algorithm is as follows:

```
Get the current word Wi, set the sentiment value Vi to 0
if(Wi belongs to Positive Word List)  set Vi=1
else if(Wi belongs to Negative Word List) set Vi=-1
else output Vi=0
if (Wi belongs to Positive Word List|| Wi belongs to
Negative Word List){
get the K words before the current word Wi
if(any word in K belongs to Privative Word List ) set Vi=
-1 *Vi
get M words before Wi and N words after Wi
if(any word in M or N belongs to Degree Word List)
set Vi=Vi* Degree
output Vi
}
```

$$v_p = \sum_{i=1}^{n} V_i$$

```
end
```

In the experiment, we randomly selected three stocks in the Growth Enterprise Market in China Toread(300005) Hanwei Electronics (300007), Huayi Brothers (300027), Their size is moderate, and the company news amount is also more suitable for this experiment.We have a Financial News crawler to get Financial News data, which is supported by Internet Financial Intelligence Laboratory, School of Information, Renmin University of China. Fig.2.is a screenshot of stock financial news text, Fig.3.shows the news number of Huayi Brothers (300027) on the timeline.In the experiment, we use the news number of stocks to represent the relative volume of information that day.

News ID	News Title	Time	Company ID	Newsbody
2010080600000007	（公司）华谊兄弟：《唐山大地震》票房达到 4.8 亿元-新闻频道-和讯网	2010-8-6	300027	全景网 8 月 6 日讯 华谊兄弟（300027, 股吧）（300027）周五盘后公告称，截至 8 月 5 日 24 时，上映 15 天，公司 2010 年重点影片《唐山大地震》的票房达到 4.8 亿元……

Fig. 2. The news store in database

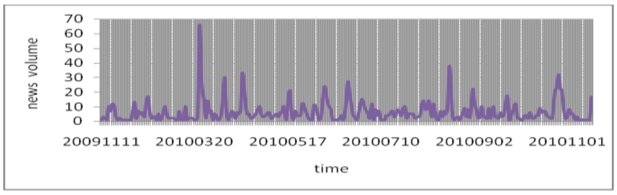

Fig.3. The news number of Huayi Brothers (300027) on the timeline

3.3 Stock price information processing

We obtained the transaction data and historical prices of three stocks from Yahoo Finance (http://finance.yahoo.com.cn),Huayi Brothers (300027)increase in shares in 2010, doubling the total share capital,we need to multiplex the right to calculate the actual price, the amendment shown in Fig.4.

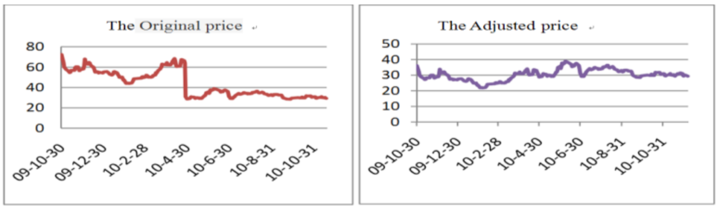

Fig.4. The original and adjusted stock price of Huayi Brothers (300027) on the timeline

The experiments selected the closing price, the opening price, the highest price and lowest price of each trading day for nearly a year for the three stocks, take the average of these four price that day the P, P = (opening price + closing price + highest price + lowest price) / 4). Financial news and stock prices are divided into a time window, we use three-day sliding window to predict the fourth day of the price. Toread (300005) a total of 370 groups (P_{t-2}, P_{t-1}, P_t, P_{t+1}) data, t = 1, 2, 3 ..., 370. Hanwei Electronics (300007) 336 groups (P_{t-2}, P_{t-1}, P_t, P_{t+1}) the data, t = 1, 2, 3 ..., 336, Huayi Brothers (300027) a total of 262 groups (P_{t-2}, P_{t-1}, P_t, P_{t+1}) data, t = 1, 2, 3 ..., 262. Use prediction models for a prediction, that is, input the relevant data during T, T-1, T-2 to forecast the stock priceT+1, and then compare the predict values to actual values, analysis the forecasting results.

3.4 Data preprocessing

The format requirements of Libsvm training data collection is : <label><index1>:<value1><index2>:<value2> ... .<label>is the target value of the training data set, for regression it could be any real number, in our experiments,<label>column is the stock price of the next time period. <index> is an integer from 1, may not be continuous. <value> is the value of predictors, or value of explanatory variables. First, we will converted the data to the format required by the above Libsvm, and use the scale Toolbox provided by Libsvm to make each factor data fall into [-1, 1] interval.

Selected training samples in accordance with the training set: test set for the ratio of 4:5, to select 4/5 of all the data as a training set. As a result, Toread (300005) got training set of 296 groups, the test set of 370 groups, Hanwei Electronics (300007) got training set of 268 groups, 336 groups for test sets, Huayi Brothers (300027) got training set of 206 groups, and the test set 262 groups.

3.5 Group experiments

Here we use the SVM regression function to exploit the relationship between financial news sentiments and stock prices of financial markets, take Libsvm as the experimental tool. Libsvm is a simple, convenient, fast and efficient SVM pattern recognition and regression package, developed and designed by LinChih-Jen of Taiwan University. There are five different types The SVM, in this experiment, we choose epsilon-SVR to do support vector regression to predict the stock price.

We designed four sets of three-day sliding window SVM prediction comparative experiment. Experiment 1: As shown in Table 1, use the stock price the data of previous three days to forecast the stock price of t+1. Experiment 2: based on Experiment 1, add the corresponding daily number of stock news as the exogenous variables. Experiment 3: based on Experiment 1, add the corresponding stock news sentimental value of the day. Experiment 4: adding corresponding daily number of stock news and stock news sentimental value at the same time based on Experiment 1.

Table 1. The input and out variables of the experiment

Indicator	**Experiment 1**	**Experiment 2**		**Experiment 3**		**Experiment 4**		
Input variables	P_{t-2}	P_{t-2}	V_{t-2}	P_{t-2}	W_{t-2}	P_{t-2}	V_{t-2}	W_{t-2}
	P_{t-1}	P_{t-1}	V_{t-1}	P_{t-1}	W_{t-1}	P_{t-1}	V_{t-1}	W_{t-1}
	P_t	P_t	V_t	P_t	W_t	P_t	V_t	W_t
			V_{t+1}		W_{t+1}		V_{t+1}	W_{t+1}
Output variable	P_{t+1}	P_{t+1}		P_{t+1}		P_{t+1}		

3.6 Calculate the evaluation score

We use the following statistical indicators to evaluate the prediction results: mean square error (MSE), the standardized mean square error (NMSE), mean absolute error (MAE), multiple correlation coefficients (SCC), and the direction of symmetry (DS), weighted direction symmetry (WDS), correct upward trend (CP), and correct downward trend (CD).

The definition and calculation formula of the statistical indicators are shown in Table 2. MSE, NMSE and MAE is an indicator which measures the deviation between the true value and the predictive value, MSE, NMSE and MAE is smaller, the predictive value is more close to the true value. The closer the multiple correlation coefficient SCC to 1, the better the forecast performance is, it can be used as an indicator to test the total effect of the forecast. DS, CP and CD are the indicators to evaluate the accuracy of the direction forecast. DS represents the proportion of all correctly predict the direction of the point, totaling 100 for all direction of the points have been correctly predicted, CP indicates the correct upward predict proportion, CD the correct downward predict proportion. So: CP + CD ⩽ DS ⩽ 100. WDS is an indicator which measures the magnitude and direction of the forecast bias. WDS punish the point of predicting the wrong direction, and reward the point of prediction in the right direction. If WDS is the smaller, the predicted direction accuracy is higher.

Table 2. The definition and calculation formula of the statistical indicators

Metrics	Calculation
MSE	$\mathrm{MSE} = \frac{1}{n}\sum (y_i - p_i)^2$
SCC	$\mathrm{SCC} = \frac{(n\sum y_i p_i - \sum y_i \sum p_i)^2}{\left(n\sum p_i{}^2 - n(\sum p_i)^2\right) \times (n\sum y_i{}^2 - n(\sum y_i)^2)}$
NMSE	$\mathrm{NMSE} = \frac{1}{n\delta^2}\sum (y_i - p_i)^2, \delta^2 = \frac{1}{n-1}\sum (y_i - \bar{y})^2$
MAE	$\mathrm{MAE} = \frac{1}{n}\sum \lvert y_i - p_i \rvert$
DS	$\mathrm{DS} = \frac{100 \times \sum d_i}{n}, d_i = \begin{cases} 1, (y_i - y_{i-1})(p_i - p_{i-1}) \geq 0 \\ 0, \text{otherwise} \end{cases}$
WDS	$\mathrm{WDS} = \frac{\sum d_i \lvert y_i - p_i \rvert}{\sum d_i{}' \lvert y_i - p_i \rvert}$ $d_i = \begin{cases} 0, (y_i - y_{i-1})(p_i - p_{i-1}) \geq 0 \\ 1, \text{otherwise} \end{cases}$ $d_i{}' = \begin{cases} 1, (y_i - y_{i-1})(p_i - p_{i-1}) \geq 0 \\ 0, \text{otherwise} \end{cases}$
CP	$\mathrm{CP} = \frac{100 \times \sum d_i}{n},$ $d_i = \begin{cases} 1, (p_i - p_{i-1}) > 0, (y_i - y_{i-1})(p_i - p_{i-1}) \geq 0 \\ 0, \text{otherwise} \end{cases}$
CD	$\mathrm{CD} = \frac{100 \times \sum d_i}{n},$ $d_i = \begin{cases} 1, (p_i - p_{i-1}) < 0, (y_i - y_{i-1})(p_i - p_{i-1}) \geq 0 \\ 0, \text{otherwise} \end{cases}$

$y_i$is the actual value, p_i is the predicted value

4 Analysis of The experimental results

In the experiment, we used four kernel functions in SVM. With different kernel functions, we also adjust the parameters respectively to get a better result, as shown in the following tables.

Table 3. Results of Toread (300005)

	Training data				Test data			
Experiment	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
MSE	0.0991	0.0759	0.0648	0.0603	0.0919	0.0761	0.2584	0.1261
SCC	0.9871	0.9901	0.9916	0.9921	0.9868	0.9890	0.9638	0.9821
MAE	0.2387	0.2205	0.2004	0.2252	0.2312	0.2184	0.2970	0.2726
NMSE	0.0129	0.0099	0.0084	0.0078	0.0132	0.0109	0.0371	0.0181
DS	65	67	71	72	63	65	68	69
CP	34	34	39	37	31	32	36	34
CD	31	32	32	35	31	33	32	34
WDS	0.4387	0.4697	0.4256	0.3282	0.4455	0.4851	0.5243	0.4292

Table 4. Results of Hanwei Electronics (300007)

	Training data				Test data			
Experiment	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
MSE	0.4800	0.3896	0.3873	0.2732	0.4198	0.3588	0.3873	0.2825
SCC	0.9553	0.9638	0.9644	0.9745	0.9614	0.9676	0.9644	0.9744
MAE	0.5224	0.4787	0.4237	0.3850	0.4832	0.4540	0.4237	0.3971
NMSE	0.0446	0.0362	0.0355	0.0254	0.0385	0.0329	0.0355	0.0259
DS	65	68	68	73	64	66	68	70
CP	35	36	36	39	34	34	36	37
CD	29	32	31	33	29	32	31	33
WDS	0.4300	0.4019	0.4210	0.3424	0.4439	0.4069	0.4210	0.4035

Table 5. Results of Huayi Brothers (300027)

	Training data				Test data			
Experiment	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
MSE	0.5661	0.5331	0.4345	0.2078	0.5066	0.4434	0.5301	0.4619
SCC	0.9633	0.9652	0.9719	0.9866	0.9591	0.9642	0.9577	0.9634
MAE	0.5596	0.5446	0.4746	0.3821	0.5314	0.5138	0.5049	0.4789
NMSE	0.0367	0.0346	0.0281	0.0135	0.0408	0.0357	0.0427	0.0372
DS	64	64	69	76	65	66	70	73
CP	31	31	34	37	32	32	34	34
CD	33	32	34	39	33	33	35	39
WDS	0.5307	0.5290	0.4588	0.3360	0.4970	0.4926	0.4731	0.4282

Table 6. Results of the average of the three stock

	Training data				Test data			
Experiment	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
MSE	0.3817	0.3329	0.2956	0.1804	0.3394	0.2928	0.3920	0.2902
SCC	0.9685	0.9730	0.9760	0.9844	0.9691	0.9736	0.9620	0.9733
MAE	0.4402	0.4146	0.3662	0.3308	0.4153	0.3954	0.4085	0.3829
NMSE	0.0314	0.0269	0.0240	0.0156	0.0308	0.0265	0.0384	0.0271
DS	64.6667	66.3333	69.3333	73.6667	64.0000	65.6667	68.6667	70.6667
CP	33.3333	33.6667	36.3333	37.6667	32.3333	32.6667	35.3333	35.0000
CD	31.0000	32.0000	32.3333	35.6667	31.0000	32.6667	32.6667	35.3333
WDS	0.4664	0.4669	0.4351	0.3355	0.4621	0.4615	0.4728	0.4203

The results data display the performance of the training set and test set in the four groups of experiments. With the increase and change in the input data, the evaluation index has more excellent performance, the accuracy of prediction has been gradually increasing, and this is more evident in the results of the training data set.

Especially from Experiment 2 to Experiment 3, We can see in the experiments of each stock, the input data change from stock price P and information volume V to stock price P and information sentimental value W, the accuracy of the forecast of the price trends DS have been greatly improved, the mean value increased from 65.7% to 68.7%, indicating that to some extent, compared to the financial information volume, financial information sentimental value is more helpful on stock price forecast.

In each stock, experiment 4 are given the best results, indicating the forecast which add information volume and financial information sentimental value based on price is more comprehensive and more accurate.

5 Discussion

In this paper, we conduct experimental research on Support Vector Machines and apply our research to stock price forecasting. Setting historical stock price data, financial information volume and information sentimental value as key factors, we got ideal forecasting result by modeling training and forecasting， analyzed and compared the results. The results indicate that information sentimental value is more influential than information volume on stock price forecasting. Comparing to the experiment with one dimensional data, the result from the experiment with both information volume and information sentimental value is more accurate.

The Internet financial information sentimental value is our major innovation on variable selection. It extends the training and forecasting model from the perspective of information, which not only uses data from the inside transaction, but also show the impact on the stock price of the outside events. The variation of variables will influence the predicted value. In our experiment, we optimized the accuracy of forecasting by minimizing the mean-square error of stock price predicted value via limiting interval.

Our paper improved the existing training and forecasting model. The learning ability of Support Vector Machines is to obtain information from data, by using time window, using more historical data to predict, it can carry more fully information. The result is better than other methods from the perspective of various predictive indicators. Thus, our experiment provides a more accurate stock price forecasting method for investors, which enable them manage investment and risk more efficiently in financial market.

But there are also some limitations in this paper, the sentimental computing is the opinion analysis of online information. The information we use the still is mainly based on news, the trend now is Micro-blogging and other social media has become a generally accepted the opinions platform, use the sentimental value of portal news to represent the overall sentimental is not comprehensive enough. In order to get more effectively and accurately predict and analysis, we need to analyze finance-related government departments, companies, social media, financial experts and commentators, even the opinions of all shareholders in the future work.

Acknowledgment

The work was supported by the Fundamental Research Funds for the Central Universities, Research Funds of Renmin University of China (10XNI029, Research on Financial Web Data Mining and Knowledge Management), and the Natural Science Foundation of China under grant 70871001.

References

1. Yaser, S.A.M., Atiya, A.F.: Introduction to Financial Forecasting. J. Appl. Intell. 6,205-213 (1996)
2. Malkiel, B.G.: A Random Walk Down Wall Street. W.W. Norton and Company Ltd, New York(1973)
3. Werner, A., Murray, Z.F.: Is All That talk Just Noise? The Information Content of Internet Stock Message Boards. J. Fin. 59, 1259-1294(2004)
4. Liang, X.: Impacts of Internet Stock News on Stock Markets based on Neural Networks. In: Wang, J. (eds.), Berlin 2005, LNCS, vol. 3497,pp. 897-903.Springer, Verlag (2005)
5. Liang, X.: Mining Associations between Web Stock News Volumes and Stock Prices. J.Int. J. Syst. Sci. 37, 919-930,(2006)
6. Sanjiv, R.D., Mike, Y.C.: Yahoo! For Amazon Sentiment Extraction from Small Talk on the Web. J. Mgt. Sci. 53, 1375-1388 (2007)
7. Wang, Y.F.: Predicting Stock Price Using Fuzzy Grey Prediction System. J. Expert Syst. Appl. 22,33-39(2002)
8. Kim, K.J., Han, I.: Genetic Algorithms Approach to Feature Discretization in Artificial Neural Networks for the Prediction of Stock Price Index.J. Expert Syst. Appl. 19,125-132(2000)

9. Cao, L.J., Tay, F.E.H.: Financial Forecasting Using Support Vector Machines.J. Neural Comput Appl.10,184-192(2001)
10. Hassan, M.R., Nath, B.: Stock Market Forecasting Using Hidden Markov Model: A New Approach. In: 5th International Conference on Intelligent Systems Design and Applications,pp.192-196. IEEE Press, Australia(2005)
11. Chang, P.C., Liu C.H.: A TSK Type Fuzzy Rule based System for Stock Price Prediction. J. Expert Syst. Appl. 34,135-144(2008)
12. Liao, Z., Wang, J.: Forecasting Model of Global Stock Index by Stochastic Time Effective Neural Network. J. Expert Syst. Appl. 37, 834-841 (2010)
13. Liang, X., Zhang, H.S., Xiao, J.G., Chen, Y.: Improving Option Price Forecasts with Neural Networks and Support Vector Regressions. J. Neural Comput Appl.72 ,3055-3065(2009)
14. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice-Hall International Inc., Englewood CliIs(1999)
15. Cao, L.J.,Tay, F.E.H.: Support Vector Machine with Adaptive Parameters in Financial Time Series Forecasting. J. IEEE Tran. on Neural Network 14, 1506-1518(2003)
16. Lee, M.C.: Using Support Vector Machine with a Hybrid Feature Selection Method to the Stock Trend Prediction. Expert Syst. Appl.36,10896-10904(2009)
17. Yeh, C.Y., Huang, C.W., Lee, S.J.: A Multiple-Kernel Support Vector Regression Approach for Stock Market Price Forecasting. J. Expert Syst. Appl.38, 2177-2186(2011)
18. Hassan, M.R., Nath, B., Kirley, M.:A Fusion Model of HMM, ANN and GA for Stock Market Forecasting. J. Expert Syst. Appl.33, 171-180(2007)
19. Bildirici, M., Erisn, O.O.: Improving Forecasts of GARCH Family Models with the Artificial Neural Networks: An Application to the Daily Returns in Istanbul Stock Exchange. J. Expert Syst. Appl. 36,7355-7362(2009)
20. Wen, Q.H., Yang, Z.H., Song, Y.X., Jia, P.F.: Automatic Stock Decision Support System Based on Box Theory and SVM Algorithm. J. Expert Syst. Appl. 37, 1015-1022(2010)
21. Zhao, Y.Y., Qin B., Liu, T.: Sentiment Analysis. J. J. Softw.21, 1834-1848(2010)
22. Chen, P.Y., Wu, S.Y., Yoon, J.S.: The Impact of Online Recommendations and Consumer Feedback on Sales. In: 25thInternational Conference on Information Systems, pp.711-724. AIS ,Washington(2004)
23. Liu, Y.: Word of Mouth for Movie: Its Dynamics and Impact on Box Office Revenue". J.J. Marketing.70, 74-89 (2006)
24. Ghose, A., Panagiotis, G.I.: Designing Novel Review Ranking Systems: Predicting the Usefulness and Impact of Reviews. In: 9th International Conference on Electronic Commerce, pp. 303-310. ACM, New York(2007)
25. Devitt, A., Ahmad, K.: Sentiment Polarity Identification In Financial News: A Cohesion based Approach. In: 45th Annual Meeting of the Association of Computational Linguistics, pp. 984-991.Association for Computational Linguistics, Prague(2007)
26. Li, N., Wu D.S.: Using Text Mining and Sentiment Analysis for Online Forums Hotspot Detection and Forecast. J. Decis. Support Syst. 48, 354-368(2010)