The two-layer SRU Neural Network Based Analysis of Time Series

Junjun Han*, Xiaoqing Zhu
School of Computer Engineering and Science, Shanghai University, Shanghai, China

*Corresponding author e-mail: hanjun5014@shu.edu.cn

Abstract. The traditional measurement model cannot meet the processing requirements of non-linear and high-noise time series data. A two-layer simple recurrent unit (DSRU) prediction model based on BPSO and SO-PMI sentiment analysis algorithm is proposed. First, the model uses the SO-PMI algorithm to calculate the stock market sentiment tendency measurement index. Then the BPSO algorithm is used to extract the characteristics of stock technical indicators with greater stock price correlation. Finally, the DSRU network model is used to predict the stock price. Taking Shanghai Pudong Development Bank stock as an example, the results show that the model has a significant improvement in prediction accuracy, reaching 65.84%, with an average absolute error of 0.0541, which proves the superiority and effectiveness of the method.

1. Introduction
With the rapidly development of the market economy, the investing in the stock market has become a hot topic in people’s daily lives. Lots of investors hope to obtain more profits by analyzing market information. Effective stock market forecasts can play a vital role in decision-making when investors choose stocks, and help investors avoid investment risks by providing deep insights into the stock market. However, due to the complexity and instability of the stock market, stock market forecasting is a very difficult but important task. Media news reports provide the market with company-related information, including the company’s news, business activities, financial conditions, and market participants’ expectations of future price changes, which can influence the order of investors’ attention to certain facts and opinions, thereby affecting investors’ investment sentiment and downward behavior, and further change the state of the market[1].

At present, most of the stock market forecasting models have the following problems: (1) Most stock forecasting models only consider stock technical indicators, and use historical stock transaction data to predict future stock market fluctuations. However, according to the latest research of behavioral finance[2], it is found that the polarity of relevant financial news and the emotional impulse of stock investors have an important impact on the stock trading market; (2) Unstructured text data of news media cannot be extracted and classified effectively; (3) There are thousands of related technical indicators calculated from stock trading data. How to effectively extract a combination of indicator features that are more relevant to stock prices is also one of the problems that need to be solved urgently. The above problems make it very difficult to predict the stock market.

In response to the above-mentioned problems, This paper proposes a new BPSO-PMI-DSRU model to forecast stock price, the model based on semantic orientation-pointwise mutual information
(SO-PMI) sentiment analysis method[3], and binary particle swarm optimization algorithm (BPSO) [4]. The model consists of a BPSO optimization algorithm for extracting technical characteristics of stocks, a SO-PMI statistical analysis algorithm for extracting emotional tendencies from network media text information, and a two-layer SRU neural network for stock price prediction. First, the model uses text analysis methods to mine the stock public opinion data, and uses the SO-PMI algorithm to determine its propensity measurement indicators. Second, the BPSO feature optimization algorithm is used to extract the most relevant feature subset of stock technical indicators, thereby effectively improving the prediction model’s performance accuracy.

2. Related work
Due to the unstructured, diversified and inexhaustible characteristics of financial news and stockholder comment data, it is a challenging task to dig out useful influence factors from massive amounts of online media information. At present, domestic and foreign scholars mostly use machine learning, text mining and other methods to study the relationship between the polarity of online media and stock market fluctuations. Wen FH et al.[5] pointed out that both positive sentiments and upward changes in sentiment have a significant positive impact on stock returns, while negative sentiments and downward changes in sentiment have no obvious impact on it. This is because the rational component plays a leading role in the market during periods of low sentiment. HK Sul et al.[6] pointed out that the emotional extreme performance of news significantly affects the volatility of the stock trading market, and analyzed the reason that some positive and negative emotional words in the news often give investors a psychological hint and induce investors to make corresponding investment decisions. Bollen et al.[7] analyzed the sentiment of stockholders on Twitter every day, and added these sentiment features to the prediction model to predict stock price. The result shows that the prediction accuracy has been greatly improved after adding news sentiment features. Li X et al.[8] used the Harvard Psychological Dictionary and Loveland-McDonald’s Financial Sentiment Dictionary to construct the emotional space, quantitatively measured news text, and projected it into the emotional space, based on the five-year historical price of the Hong Kong Stock Exchange Build predictive models with relevant news texts. Dang et al.[9] proposed a model that uses financial news and sentiment dictionaries to predict stock price, and uses the Stock2Vec embedding model to classify news sentiment, and combines stock transaction data to establish a dual-stream gated recurrent unit (TGRU) for stock market prediction. Xu Y et al.[10] proposed a hybrid forecasting model combining k-means clustering and ensemble learning based on stock technical indicators and news information. Experimental results on a number of Chinese financial stocks show that the hybrid forecasting model obtains the best stock price forecast accuracy.

3. Method design
In this paper, the SO-PMI statistical analysis algorithm is used to qualitatively analyze the sentiment value of news and stock comment data, and the BPSO feature optimization algorithm is used to obtain the optimal feature subset of stock technical indicators, which is input to the two-layer simple recurrent unit (DSRU) network model in the process of predicting stock prices, the overall method flow is shown in Figure 1.

Firstly, use crawler technology to obtain news and stock comment text sequences, and use python stammering word segmentation tool[11] to extract topic keywords, combined with HowNet Sentiment Dictionary[12], through SO-PMI emotional word tendency analysis method calculates the mutual information between the target word and the words in the basic sentiment dictionary, to extract the news polarity and investor sentiment tendency characteristic index. Secondly, the BPSO feature extraction algorithm is used to extract the feature subset of stock price technical indicators with the greatest correlation to stock price fluctuations.

For the part of the network prediction model, due to the time-series correlation between the input features and the stock price, a two-layer simple recurring unit (DSRU) network model is used to predict the stock price trend.
3.1. **SO-PMI Sentiment Tendency Analysis**

3.1.1. *News and stock data acquisition.* This article selects Shanghai Pudong Development Bank stock transaction data from October 10, 2019 to October 10, 2020 as the research object. Using crawler technology to obtain relevant news and stock comment information from "Sina Finance Network" and "Stock Bar Forum in Oriental Fortune Network", including the time, title and content of the news and comment. Use python to stutter word segmentation, and get stock-related subject words according to regular expressions.

3.1.2. *Sentiment analysis.* First of all, using HowNet as the basic sentiment dictionary, by calculating the PMI value of the mutual information between the target words and the words in the HowNet sentiment dictionary, introducing the sentiment orientation (OS) of the calculated words to determine the sentiment orientation of the target word. The formula for calculating the PMI value of two words is as follows:

$$PMI(word_1, word_2) = \ln\left(\frac{P(word_1 + word_2)}{P(word_1) \times P(word_2)}\right)$$

Where $P(word_1 + word_2)$ represents the probability that the words word_1 and word_2 appear together, $P(word_1)$ and $P(word_2)$ represent the probability of the two words appearing separately, and $PMI(word_1, word_2)$ represents the point between the two words mutual information, the result indicates: when PMI > 0, two words are related; when PMI = 0, two words are independent of each other; when PMI < 0, two words are not related.

Let $W=\{word_1, word_2, ..., word_n\}$ be the set of all thematic keywords contained in a certain news or stock review, and take the "positive emotional word set" and "negative emotional word set" in the HowNet sentiment analysis vocabulary set as optimistic. The standard word Pword and the bearish standard word Nword. The bullish standard word means that the stock price may rise, and the bearish standard word means that the price is falling. Calculate the mutual information between each word in the W set and Pword and Nword separately. The emotional tendency of the word can be judged by the difference of the mutual information between the two points. The calculation formula is as follows:

$$PMI(Pword) = \sum_{word_1 \in Pword} PMI(word_1, P)$$

$$PMI(Nword) = \sum_{word_1 \in Nword} PMI(word_1, N)$$
\[ SO\text{-}PMI(word_i) = PMI(Pword) - PMI(Nword) \]  \hspace{1cm} (4)

Where \( \text{word}_i \in W, i=1,2,...,n \). Calculate the average of the SO-PMI of all words in the set \( W \), use 0 as the threshold, and the calculated value includes three states:

\[
\overline{SO\text{-}PMI} = \begin{cases} 
> 0; & \text{Positive tendency} \\
= 0; & \text{Neutral orientation} \\
< 0; & \text{Negative tendency}
\end{cases} \hspace{1cm} (5)
\]

Using the above methods to analyze the tendency of news and stock review texts, you can obtain the emotional polarity characteristics of news and stock reviews, which are optimistic about the use of "1" means, "0" means neutral, and "-1" means bearish. Part of the processed data is shown in Table 1.

| Date       | Title                                                                 | Key Words                        | Propensity Value |
|------------|----------------------------------------------------------------------|----------------------------------|------------------|
| 2019-10-10 | At the beginning of the market, there is a strong signal, and the main capital strikes 11 shares | Market, Strong signal            | 1                |
| 2019-10-11 | The big wave just started to fall, so hurry up and run for clearance, and then consider coming back a year later | Fall, Hurry up, Clearance        | -1               |
| 2020-10-10 | Domestic stocks are strong and debt is weak, market expectations are optimistic, and risk appetite has rebounded significantly | Strong, Weak, Optimistic, Rebounded significantly | 1                |

3.2. BPSO Feature Extraction Algorithm

The PSO algorithm is a particle swarm optimization algorithm proposed by Kennedy et al.[13], which can only be used to solve continuous problems. For discrete problems, Kennedy et al. discretized the position information in the algorithm into binary 0 or 1, and obtained BPSO discrete optimization. The speed and position update formula is as follows:

\[
v_{id}^{k+1} = \omega \cdot v_{id}^k + c_1 r_1 \cdot (P_{id} - x_{id}^k) + c_2 r_2 \cdot (P_{gd} - x_{id}^k) \hspace{1cm} (6)
\]

\[
x_{id}^{k+1} = \begin{cases} 
1, & \xi < S(v_{id}) \\
0, & \xi \geq S(v_{id})
\end{cases} \hspace{1cm} (7)
\]

Where \( \omega > 0, c_1, c_2 \) are learning factors, \( r_1, r_2 \) are random numbers between \([0, 1]\), \( v_{id} \) and \( x_{id} \) represent the velocity and position of particle \( i \) in the \( d \) dimension, and \( k \) represents iteration the number of times, \( P_{id} \) is the historical optimal value, \( P_{gd} \) is the global optimal value, \( S(v_{id}) \) is a sigmoid function that maps the speed to \((0 \sim 1)\).

3.3. SRU Network Model

Lei T et al.[14] proposed a variant of RNN deep learning model SRU in 2017. This model adjusts the structure of the existing gating unit, removes the dependency on the calculation at the front and back, and simplifies the process of state calculation. It can achieve the same parallelism as CNN and feedforward neural network. The SRU unit structure is shown in Figure 2.
The basic form of the SRU structure includes a forget gate $f_t$ and a reset gate $r_t$. For the input $x_t$ at a given time $t$, the calculation formula is as follows:

$$\tilde{x}_t = Wx$$  \hspace{1cm} (8)

$$f_t = \sigma_f(W_f x_t + b_f)$$  \hspace{1cm} (9)

$$C_t = f_t \odot C_{t-1} + (1 - f_t) \odot x_t$$  \hspace{1cm} (10)

$$r_t = \sigma_r(W_r x_t + b_r)$$  \hspace{1cm} (11)

$$h_t = r_t \odot \tanh(C_t) + (1 - r_t) \odot x_t$$  \hspace{1cm} (12)

Where $W$, $W_f$, $b_f$ are the parameter matrices to be trained, $\sigma_f$ and $\sigma_r$ are the sigmoid activation functions, tanh is the hyperbolic tangent activation function of the hidden state, and the forgetting gate $f_t$ is used to adjust the internal state $C_t$. The reset gate $r_t$ is used to combine the data input signal $x_t$ and the current cell state $C_t$ to calculate the output state $h_t$ of the model.

### 3.4. Stock prediction network model establishment

Figure 3 shows a stock prediction model based on a two-layer SRU network structure, where $x_1, x_2, \ldots, x_t$ are input features after preprocessing, and each item contains a data set that affects stock price on a certain day, and then Use the BPSO optimization algorithm for feature extraction to obtain the optimal feature subset, which is input to the first layer of SRU unit, and the output result is processed by the update gate and reset gate, and then input to the next layer of SRU unit to obtain the output vector $y$, and finally through A fully connected layer is used to obtain the predicted value $y_{pre}$. Where $h^{(1)}$ and $h^{(2)}$ are the hidden states of the first and second layer SRU, respectively.
4. Experiment analysis

4.1. Stock price prediction experiment

4.1.1. Experimental data. The stock transaction data selects the historical stock transaction data of Shanghai Pudong Development Bank. The data includes basic characteristic indicators such as transaction date, closing price, opening price, and highest price. There are about 30 kinds of technical indicators calculated through basic stock data. This article selects the most influential feature subsets through the BPSO feature selection model, including: moving averages of similarities and differences (MACD, DIF, DEA), energy tides (OBV), rate of change index (ROC), stochastic index (KDJ) and bollinger index (BOLL) and other 14 stock technical indicators, all of which are calculated from basic stock trading data. Part of the technical index data is shown in Table 3 below.

Table 2. Stock technical indicators characteristic data.

| Technical index | 2019-10-10  | 2019-10-11  | ...    | 2020-10-10  |
|-----------------|-------------|-------------|--------|-------------|
| MACD            | 0.007122    | -0.003892   | ...    | 0.086518    |
| EDA             | 0.043128    | 0.042977    | ...    | 0.050697    |
| DIF             | 0.046689    | 0.041031    | ...    | 0.093956    |
| K               | 67.925317   | 47.946077   | ...    | 55.741372   |

Because different stock price technical features have different dimensions and dimensional units, the value range of each dimension feature component is quite different, and the difference between the feature components can easily cause the feature selection model to become complicated and reduce the efficiency of feature selection. Therefore, normalizing the input feature variables can solve the problem of comparability between the technical features of stock prices and reduce the error output. The normalized value range of the sample data selected in this article is [-1, 1], and the normalized data formula is as follows:

\[
\tilde{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{12}
\]

Where \(x_{\min}\) is the minimum value of each feature component in the sample data, and \(x_{\max}\) is the maximum value of each feature component in the sample data.
4.1.2. Evaluation index selection. In order to verify the effect of the two-layer SRU stock price prediction model that integrates stock public opinion into the stock technical characteristics, the experiment uses accuracy (P), average absolute error (MAE), mean square error (MSE) and average absolute percentage error (MAPE) as the result evaluation index of this experiment. The calculation formula is:

\[
P = \frac{A}{N} \times 100\% 
\]

(13)

\[
\text{MAE} = \frac{1}{N} \left( \sum_{i=1}^{N} |y - \hat{y}| \right) 
\]

(14)

\[
\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y - \hat{y})^2 
\]

(15)

\[
\text{MSPE} = \frac{\sum_{i=1}^{N} \left| \frac{y - \hat{y}}{y} \right|}{N} \times 100\% 
\]

(16)

Where A represents the number of samples for which the stock price rises and falls are judged correctly, N represents the total number of samples, y represents the actual value of the stock price, \( \hat{y} \) represents the predicted value of the stock price, and P represents the accuracy rate.

4.2. Results and analysis of stock price prediction experiment

4.2.1. Comparison results and analysis before and after integrating into the stock market public opinion. In order to verify that the public sentiment of the stock market has a certain influence on the stock price, the BPSO-PMI-DSRU model of stock market public sentiment combined with stock technical indicators and the BPSO-DSRU model of single stock technical indicators are compared for experiments. It can be concluded from table 4 that the MAE, MSE and MAPE have all been significantly reduced, while the prediction accuracy rate has been greatly improved, which shows that the inclusion of stock market public opinion can improve the accuracy of the prediction model.

**Table 3. Comparison of experimental results.**

| Method      | P  | MAE  | MSE  | MAPE   |
|-------------|----|------|------|--------|
| BPSO-DSRU   | 56.37 | 0.0545 | 0.0025 | 2.4023 |
| BPSO-PMI-DSRU | 65.84 | 0.0541 | 0.0016 | 2.2998 |
Figure 4. Comparison of forecast value and actual closing price.

Figure 4 shows the predicted value and true value of the stock closing price on the test set of the BPSO-PMI-DSRU and BPSO-DSRU models. It can be intuitively seen from the figure that the predicted value of BPSO-PMI-DSRU is closer to the closing price. The real value, therefore, the prediction of the stock market public opinion into the technical indicators can have better results than the single technical indicator prediction.

4.2.2. Comparison results and analysis with other methods. In order to verify the superiority of the model proposed in this article, the method of this article is compared with other network models, including support vector machine (SVM), BP neural network model, long and short-term memory neural network (LSTM) and gated recurrent unit (GRU). The sentiment characteristics of stock market public opinion combined with stock technical indicators are input into the network model for stock price prediction. The experimental results of the method in this paper and other methods are shown in Table 5. Through the analysis results, it can be seen that the BPSO-PMI-DSRU model proposed in this paper has achieved better performance in accuracy and MAE, MSE and MAPE indicators, which verifies that the model has the best effect in stock price prediction.

Table 4. Comparison of experimental results with different methods.

| Method            | P    | MAE  | MSE  | MAPE  |
|-------------------|------|------|------|-------|
| BPSO-DSRU         | 56.37| 0.0545| 0.0025| 2.4023|
| BPSO-PMI-DSRU     | 65.84| 0.0541| 0.0016| 2.2998|
| BPSO-PMI-LSTM     | 65.27| 0.0583| 0.0019| 2.4023|
| BPSO-PMI-GRU      | 64.34| 0.0591| 0.0027| 2.5798|
| BPSO-PMI-DSRU     | 65.84| 0.0541| 0.0016| 2.2998|

5. Conclusion
This paper proposes a new BPSO-PMI-DSRU model, which combines stock market public opinion and stock transaction historical data in media effectively predict stock prices. A large number of comparative experiments have verified the effectiveness and superiority of the method. The next step is to further expand the source of financial texts, avoid the problem that the uneven distribution of individual stock-related news and comment data affects the prediction effect, and further explore whether the quantitative analysis of the stock market public opinion can improve the accuracy of the stock price prediction model.
References

[1] Hagenau M, Liebmann M, Nedwig M, et al. (2020) Automated news reading: stock price prediction based on financial news using context-capturing features [J]. Decision Support System, 55(3):685-679.

[2] Daniel Kahneman, Amos Tversky. (2014) Prospective Theory: An analysis of decision under risk. 8 (29-30):29-124.

[3] Turney P D, Littman M L. (2003) Measuring Praise and Criticism: Inference of Semantic Orientation from Association [J]. ACM Transactions on Information Systems, 21(4):315-346.

[4] Liao TW, Kuo R J. (2018) Five discrete symbiotic organisms search algorithms for simultaneous optimization of feature subset and neighborhood size of KNN classification models [J]. Applied Soft Computing, 64:581-595.

[5] Wen FH, Xiao JL, Huang CX, Chen XH. (2014) Research on the influence of investor sentiment characteristics on stock price behavior [J]. Journal of Management Science, 17(03): 60-69.

[6] HK Sul, AR Dennis, Yuan L. (2017) Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns [J]. Decision Sciences, 48(3).

[7] Bollen J, Mao H, Zeng X. (2011) Twitter mood predicts the stock market [J]. Journal of Computational Science, 2(1):1-8.

[8] Li X, Xie H, Chen L, et al. (2014) News impact on stock price return via sentiment analysis [J]. Knowledge Based Systems, 69(Oct.):14-23.

[9] Dang LM, Sadeghi-Niaraki A, Huynh HD, et al. (2018) Deep Learning Approach for Short-Term Stock Trends Prediction Based on Two-Stream Gated Recurrent Unit Network [J]. IEEE Access, 6, 55392-55404.

[10] Xu Y, Yang C, Peng S, Nojima Y. (2020) A hybrid two-stage financial stock forecasting algorithm based on clustering and ensemble learning [J]. Applied Intelligence, 1-16.

[11] Sun J. (2012) Jieba chinese word segmentation tool [J]. (2018-01-21)(2018-06-25). https://github.com/fxsjy/jieba.

[12] HowNet[EB/OL]. http://www.keenage.com. 2007.

[13] Kennedy J, Eberhart RC. Particle swarm optimization [A]. (1995) Proceedings of International Conference on Neural Networks IV [C]. Piscataway NJ: IEEE Service Center, 1942-1948.

[14] Lei T, Zhang Y, Wang SI, et al. (2018) Simple Recurrent Units for Highly Parallelizable Recurrence [C]. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels: ACL, 4470−4481.