Research on the Improvement of Big Data Feature Investment Analysis Algorithm for Abnormal Trading in the Financial Securities Market

Jie Zou[1*], Wenkai Gong[1], Guilin Huang[1], Gebiao Hu[1], Wenbin Gong[2]

[1]State Grid Jiangxi Electric Power Company Construction Branch,
Nanchang 330000
China

[2]Jiangxi Power Transmission and Transformation Construction Company
Nanchang 330000
China

*Corresponding author: zoujoeok@126.com

Abstract—Traditional investment analysis algorithms usually only analyze the similarity between financial time series and financial data, which leads to inaccurate and inefficient analysis of investment characteristics. In addition, the trading volume of financial securities market is huge, the amount of investment data is also very large, and the detection of abnormal transactions is difficult. The aim of feature extraction is to obtain mathematical features that can be recognized by machine. Different from the traditional methods, this paper studies and improves the big data investment analysis algorithm of abnormal transactions in financial securities market. After processing the captured trading data of financial securities market, the big data feature of abnormal trading is extracted. Combined with the abnormal trading and the financial securities market, the investment strategy is determined. The optimization objective function is set and the genetic algorithm is used to improve the investment analysis algorithm. The simulation experiment verifies the improved investment analysis algorithm, and the average Accuracy of investment analysis is increased by at least 11.24%, the ROI is significantly improved, and the efficiency is higher, which indicates that the proposed algorithm has ideal application performance.

Keywords—financial securities market; abnormal trading; big data feature; investment analysis algorithm; algorithm improvement.

I. INTRODUCTION

As the development of economy gets better and better, people's income is also rising. The traditional bank savings have been unable to meet the needs of most people's financial investment. More and more people put their funds into the financial securities market. The financial securities market is a place where all kinds of securities are issued and exchanged, including stocks and bonds. With the continuous accumulation of liquidity in the financial securities market, many illegal acts are also born. In our country, the CSRC is relied on to supervise the abnormal trading behavior of the financial securities market, so as to ensure the fair operation of the financial securities market and safeguard the legitimate rights and interests of investors [1]. Mathematical finance is a branch of finance system.

Investment advisors help clients design and manage their portfolios by understanding their investment and financial needs, their risk appetite, and their financial knowledge. Artificial Intelligence (AI), with its enhanced "camera decision-making" capabilities, not only provides every customer with differentiated investment strategies, but also has the potential to greatly improve the quality of service by going beyond the developer's cognitive model. AI customer service has rapidly become a new business growth point for financial investment services. In the era of big data, "data is power", people can solve problems directly based on big data, not on the premise of transforming data into knowledge. In view of the huge amount of transaction data in the capital market, it is meaningful to use the big data framework to find rules and make judgments from the transaction data to support the investment portfolio decision-making of investors, which is also the source of the core competitiveness of the business.

In recent years, related theories of finance have been developing with the emergence of portfolio and other models. In the past related research, there are still more investment problems. Mean-variance portfolio is mainly

used in the following two situations, namely, the investment behavior requires the market environment to consider the purchase of products as simply as possible. However, with the gradual diversification of people's consumption behaviors, investors are unable to keep a rational mind when making investments and keep calm in the face of risks. Therefore, such a form of investment portfolio is no longer applicable to the modern market. It also makes the relevant researchers continue to optimize and upgrade the investment theory, the theoretical system of comprehensive analysis of the problems and take corresponding measures to improve and implement the related to portfolio theory and the effective expansion of the securities investment behavior research contents, mainly on how to optimize and improve the content of the portfolio theory.

However, with the continuous increase of trading volume in the financial securities market, in order to obtain more benefits, the issuers in the financial securities market will disturb the order of the financial securities market through a variety of abnormal trading behaviors. In addition, the non-standard operation of some listed companies will lead to the lack of investment function of the securities market. These abnormal trading behaviors in the financial securities market not only seriously disturb the normal operation order of the market, but also seriously damage the return of ordinary investors in the financial securities market.

Literature [2] puts forward the optimal investment view under the condition of variable adjustment and minimum insurance. Generally speaking, there are small batches of operations before abnormal transactions occur in financial securities market, so this method cannot analyze the data of transactions. Literature [3] proposed an investment analysis method based on empirical analysis of real options. This algorithm requires a large number of prior data, and its processing and analysis efficiency is relatively low, so there are certain limitations when it is used. Literature [4] proposes to select "Internet finance" and "big data" concept stocks based on the limited attention theory in behavioral finance, design corresponding matching stocks with the help of propensity score method, and compare and test the causal correlation model between investors' attention and stock market trading volume and yield. Through empirical analysis, it is found that: (1) investors' attention significantly affects concept stocks rather than matching stocks, and has a "plate phenomenon"; (2) Investors' attention has a strong causal relationship with the trading volume of concept stocks, and is related to the return of concept stocks at the significant level of 10%; (3) The relationship between investors' concern and the trading volume of concept stocks is asymmetric, but it does not exist with the yield of concept stocks. This study uses big data to analyze the behavior characteristics of investors, but can not screen out abnormal transactions.

At present, the amount of trading data in the financial securities market is increasing, but the traditional investment analysis algorithm can not meet the needs of the current investment managers. Combined with the nature and feature of trading data in the financial securities market, this paper uses big data analysis technology to extract the data feature of abnormal trading behavior in the financial securities market, so as to optimize and improve the investment analysis algorithm. Therefore, based on the above analysis, in order to improve the reliability of the investment analysis algorithm, this paper will study the improvement of big data feature investment analysis algorithm for abnormal trading in the financial securities market.

II. RESEARCH ON THE IMPROVEMENT OF BIG DATA FEATURE INVESTMENT ANALYSIS ALGORITHM FOR ABNORMAL TRADING IN THE FINANCIAL SECURITIES MARKET

A. Capture abnormal trading big data in the financial securities market

In the process of capturing trading data in the financial securities market, due to a variety of reasons, market trading data will appear missing, inconsistent and other data, affecting abnormal data detection and capture. In order to eliminate the interference of noise data on anomaly detection and capture results, the following steps are used to clean the market transaction data. For data type error, coding error, format error and other forms of direct data error, the cause of the error can be directly corrected. In addition, in order to avoid the impact of inconsistent data specifications on the accuracy of feature extraction of abnormal trading big data in the financial securities market. Through data cleaning, financial securities market trading data specifications are unified and content is abstracted. That is, the same securities, stocks and other trading behavior data are made to have the same data object name and type, and the data value unit, format and length are consistent. If a certain type of securities and stocks in the financial securities market are in the period of closing adjustment, the transaction data is empty, and the data is filled in the form of 0 code. After data filling, the data is compressed on the premise of ensuring data integrity. Data compression can effectively reduce the amount of data and improve the processing, storage and transmission rate of subsequent data.

The trading data of the financial securities market contains multiple categories, and the features of different categories of data are different. The data features are coded by hot coding. N-bit 0 and 1 binary codes are used to represent the state of scheduling data. If each feature of the trading data in the financial securities market corresponds to possible feature labels with the number of n, the data features are transformed into n-dimensional binary feature sequences after the exclusive hot coding [5].

Because the trading data of financial securities market is dynamic and will change constantly. Therefore, it is necessary to preprocess the network big data stream. Window technology is simple and easy to understand, and it usually takes the latest data as the main data, which meets the requirements of capturing abnormal trading big data stream in the financial securities market. When the new data arrive, the oldest data in the window will be deleted from the window and the new data will enter the window. By keeping the amount of data in the window fixed, the data in the current window can be processed and the processing efficiency can be improved. In view of the

higher efficiency of the sliding window processing method, this paper chooses the sliding window method to process the data flow to be detected in the network [6].

Assuming that the trading data flow of financial securities market is $DS = d_{t+1}, \cdots, d_{t+i}, \cdots, d_{t+n}, \cdots$, the sliding window size is

$$d(z_g, x_i) = \frac{\sum_{i=1}^{g}(z_g - x_i)(z_g' - x_i)}{\sqrt{\sum_{i=1}^{g}|z_g - x_i|^2} \sqrt{\sum_{i=1}^{g}|z_g' - x_i|^2}}$$, the sampling size

is z_g, and the data flow is divided into data blocks of the same size. The data in the sliding window is given a weight z_g' which is a time-dependent function. The time t of the weight $d(z_g, x_i)$ is the residence time of each data point in the sliding window. The size of the sliding window is determined by the number of data blocks, and the sampling size can control the block size of the data stream. The basic size of each window is $\left[N/K\right]$ ($\left[\ \right]$ represents the rounding in the negative direction of data flow direction) [7].

The Figure 1 shows the schematic diagram of using the function sliding window to process the trading data flow of the financial securities market.
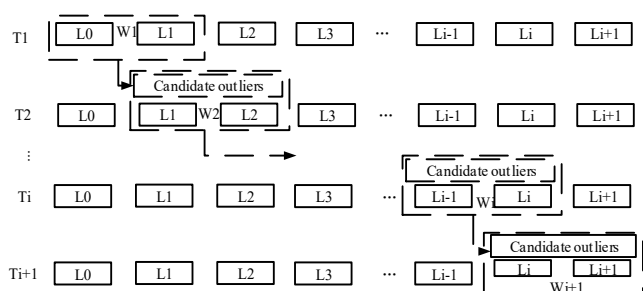

Figure 1. Schematic diagram of sliding window processing network big data flow

When the sliding window is used to process the trading data stream of the financial securities market, the average value of the data in the window is calculated according to the time sequence of the data stream. The difference between the average value and the average value of all the data on the normal trading day of financial securities market is made. If the difference is greater than the previous difference, the data point is regarded as a candidate outlier. Due to the difference of data point distribution feature in different time periods of data stream, when a data point is determined as a candidate abnormal data point in a window, the data point may be normal when the window slides to the next time [8]. Therefore, only when the point is judged as a candidate abnormal data point in the subsequent multiple sliding windows, can the point be regarded as the abnormal data point to be tested. After capturing abnormal trading data stream of the financial securities market by using sliding window technology, the features of data stream are extracted.

B. Extract the features of abnormal trading big data in the financial securities market

Due to the real-time, heterogeneous and large amount of data, it is more difficult to extract abnormal behavior data features from the trading data of the financial securities market. In order to improve the efficiency of investment analysis algorithm, this paper uses machine learning principle to extract the features of abnormal trading big data in the financial securities market.

Machine learning reduces the difficulty of trading big data feature processing by reducing the dimension of trading data in the financial securities market. First, the probability distribution of abnormal trading big data in high-dimensional space in the financial securities market is calculated, and the calculation formula is as follows [9].

$$p_{j|i} = \frac{\exp\left(-\|x_i - x_j\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq i}\exp\left(-\|x_i - x_k\|^2 / 2\sigma_i^2\right)} \tag{1}$$

In the formula (1), x_i and x_j are any two data in the big data of abnormal trading in the financial securities market. x_k is another trading data selected randomly. σ_i is the Euclidean distance between the two abnormal trading data x_i and x_j in the financial securities market. According to the above calculation results, the probability distribution of data in high-dimensional space can be obtained, so as to measure the similarity of data in the whole data. Similarly, according to the principle shown in the formula (1), the same data probability distribution is constructed in low dimensional space, and the calculation formula is as follows [10].

$$q_{j|i} = \frac{\exp\left(-\|y_i - y_j\|^2\right)}{\sum_{k \neq i}\exp\left(-\|y_i - y_k\|^2\right)} \tag{2}$$

In the formula (2), y_i and y_j are the two trading data corresponding to the low dimension space and the high dimension space. y_k is the data selected in the low dimension space. When calculating the probability distribution of data in high and low dimensional space, the final probability distribution value should be as close as possible. Then, the probability distribution divergence is calculated according to the following formula to determine the feature difference between abnormal trading data and normal trading data.

The calculation formula of probability distribution divergence is as follows.

$$\frac{\partial C}{\partial y_i} = \frac{4\sum_{j}(p_{j|i} - q_{j|i})(y_i - y_j)}{\left(1 + \|y_i - y_j\|^2\right)} \tag{3}$$

According to the probability distribution divergence of data in two spatial dimensions, the distance between all data in two spatial dimensions is calculated by clustering algorithm, and the calculation formula is as follows [11].

$$d\left(z_g, x_i\right) = \frac{\sum_{i=1}^{g}\left(z_g - x_i\right)\left(z_g{}' - x_i\right)}{\sqrt{\sum_{i=1}^{g}\left|z_g - x_i\right|^2}\sqrt{\sum_{i=1}^{g}\left|z_g{}' - x_i\right|^2}} \qquad (4)$$

In the formula (4), z_g is the cluster center in high-dimensional space. $z_g{}'$ is the cluster center in low dimensional space, that is, the adjacent sample to the cluster center. x_i is the feature vector of abnormal trading data. The cluster centers are constantly changed. The distance between the cluster center and the sample to be clustered is calculated iteratively until the calculated clustering distance is constant, and the cluster stops. The larger the clustering distance $d\left(z_g, x_i\right)$ is, the lower the similarity between the features of the trading data to be clustered and the feature vector of the cluster center is. The probability distribution divergence of financial trading data in two spatial dimensions is taken as the threshold of clustering algorithm. If the clustering distance is greater than the threshold, the data feature is the abnormal trading data feature of financial securities market. After extracting the features of abnormal trading big data in the financial securities market, combined with the common investment methods, the investment strategy under abnormal trading behavior in the financial securities market is determined.

C. Determine the investment strategy under abnormal trading behavior in the financial securities market

At present, the commonly used investment strategies in the financial securities market can be divided into benchmark strategy, chasing-up strategy, chasing-down strategy, pattern matching strategy and meta learning strategy. This section will analyze the specific use of these five strategies to determine how to invest in the case of abnormal trading behavior in the financial securities market.

Benchmark strategy mainly refers to an investment strategy in which investors do not intervene too much. Generally, it can be divided into purchase holding strategy and fixed proportion strategy. The purchase holding strategy refers to that after a certain amount of initial capital is invested in the financial securities market, the proportion of capital investment is no longer adjusted according to the operational risk of the market. This kind of investment strategy is an ideal post event strategy which is not suitable for large amount of investment in the case of abnormal trading behavior. Fixed proportion strategy is to adjust the proportion of investment funds in different risk assets in each fixed trading period, so as to obtain the cumulative rate of return. However, this kind of investment method often results in higher operating trading costs [12]. Therefore, the investment proportion can only be adjusted in a specific trading period.

Chasing-up strategy is to increase the investment proportion of risk assets with better performance in each investment trading period, in order to obtain higher investment returns. This kind of investment strategy is risky and easy to cause the investment funds to be locked up. The chasing-down strategy is a kind of investment method

which can appropriately increase the proportion and weight of the risk assets when it is not good in the current investment trading period. But it is predicted to be good in the future growth trend. This kind of investment strategy requires high investment ability of investors. Pattern matching strategy needs more accurate prediction of the future financial securities market. Generally, this kind of investment strategy has a longer investment period and higher requirements for investors' investment psychology [13]. Meta learning strategy is to use the learning algorithm of artificial intelligence, which needs a lot of prior knowledge and is not suitable for ordinary investors.

When there is abnormal trading behavior in the financial securities market, benchmark strategy, chasing-down strategy and pattern matching strategy should be combined and used as the main strategies. Catch-up strategy is used as the supplemental strategy to invest. According to the investment strategy in the case of abnormal trading behavior in the financial securities market, the investment analysis algorithm is improved.

D. Complete the improvement of investment analysis algorithm in the financial securities market

The essence of investors' investment behavior in the financial securities market is to hope that the net present value of the risk assets they choose will increase, so as to improve the return on investment. Therefore, this paper takes the maximum net present value as the objective of investment analysis algorithm improvement and optimization, and establishes the following optimal objective function [14].

$$\max NPV = \sum_{t=1}^{n} \frac{C_t}{\left(1+k\right)^t} - I \qquad (5)$$

In the formula (5), n is the investment period set by the investor for the current purchase of risky assets. k is the minimum capital return rate required for the investor when selecting the investment strategy. C_t is the capital return of the investor within the set investment period. I is the investment expenditure. Genetic algorithm is used to optimize the above objective function solving algorithm, so as to improve the investment analysis algorithm.

According to the process of biological evolution, individuals in the population are selected according to the size of adaptability, and the probability of individuals being selected can be calculated by the following formula [15-17].

$$p_s\left(x_j\right) = \frac{f\left(x_j\right)}{\sum_{i=1}^{n} f\left(x_j\right)}, j = 1, 2, \cdots, n \qquad (6)$$

In the formula (6), $f\left(x_j\right)$ is the individual fitness value. The higher the fitness value is, the more likely the individual will be selected. The selected individuals are arranged according to the fitness value, and the selected individuals are crossed and mutated. A random number is selected in the [0, 1] interval. If the random number is less than the individual selection probability, the individual is selected as the crossover parent, and then the parent is randomly paired according to the following formula.

$$\begin{cases} v_1' = cv_1 + (1-c)v_2 \\ v_2' = cv_2 + (1-c)v_1 \end{cases} \qquad (7)$$

In the above formula, c is a random number on the interval [0, 1]. v_1 and v_2 are the two cross variant parents. v_1' and v_2' are offspring after cross mutation. When the fitness function value of the genetic operator no longer changes, the iteration is stopped, and the improvement of the investment analysis algorithm based on genetic algorithm is completed. Through the above content, this paper completes the research on the improvement of big data feature investment analysis algorithm for abnormal trading in the financial securities market. The performance of the algorithm will be tested as follows [18-20].

III. SIMULATION EXPERIMENT AND ANALYSIS

This paper studies the improvement of big data feature investment analysis algorithm for abnormal trading in the financial securities market. This section will verify the improvement effect of the above algorithm through simulation experiments.

A. Experimental environment and data processing

This paper selects the historical trading data stored in a financial securities market as the research object, and selects the trading data of all trading days in the financial securities market from March 14, 2017 to March 14, 2018 from the trading data storage terminal of the financial securities market as the test data of the simulation experiment.

The experiment is carried out on the simulation platform with Intel Core i7, CPU at 3.89 GHZ, 16 G memory and 500G hard disk storage. The experimental simulation platform is equipped with professional securities simulation operation software, which is used to verify the investment analysis algorithm of the financial securities market.

B. Experimental process

This experiment uses the traditional investment analysis algorithm mentioned in literature [2] and the traditional investment analysis algorithm mentioned in literature [3] to compare with the improved big data feature investment analysis algorithm of abnormal trading in the financial securities market to complete the simulation example verification. In Tongdaxin Securities simulation software, three investment analysis algorithms are used to make investment. By comparing the rate of return on investment and Accuracy of investment analysis, the practical application effect of the three algorithms is compared.

C. Experimental results

The actual data of three investment analysis algorithms are shown in the table below. By analyzing the data in the table, the final conclusion of this experiment is obtained.

Table 1. Comparison of simulation application verification data of three investment analysis methods

Investment behavior number	Literature [2] algorithm		Literature [3] algorithm		Algorithm in this paper	
	Return on investment/%	Accuracy of investment analysis/%	Return on investment/%	Accuracy of investment analysis/%	Return on investment/%	Accuracy of investment analysis/%
1	0.186	59.4	0.328	63.7	0.985	75.7
2	0.206	59.6	0.513	66.2	0.898	78.4
3	0.229	60.2	0.377	64.4	0.977	79.5
4	0.209	58.7	0.332	65.6	0.911	74.6
5	0.222	57.3	0.449	65.8	0.916	76.9
6	0.352	60.8	0.514	67.0	0.969	75.2
7	0.184	59.1	0.359	65.5	0.963	79.6
8	0.214	58.7	0.469	65.3	0.944	77.3
9	0.215	61.1	0.454	67.4	0.996	76.1
10	0.202	58.6	0.436	65.8	1.006	75.8

It can be seen from the above table that when the algorithm mentioned in the two pieces of literature is used for investment, the return on investment is far lower than that of the method in this paper. The average Accuracy of investment analysis of the algorithm in reference [2] is 59.35%, the average Accuracy of investment analysis of the algorithm in reference [3] is 65.67%, and the average Accuracy of investment analysis of the algorithm in this paper is 76.91%, which is higher than the other two algorithms. Combined with the Accuracy of investment analysis of the algorithm and the data change, the practical application effect of this method is better. To sum up, the big data feature investment analysis algorithm for abnormal trading in the financial securities market studied in this paper has good data processing effect, can at least improve the Accuracy of investment analysis of about 11.24%, and effectively improve the return on investment.

On the basis of the above experimental results, the algorithm in reference [2] and the algorithm in reference [3] are used as comparisons to compare the analysis efficiency of the proposed algorithm with the characteristic investment analysis efficiency of the big data of abnormal transactions in the financial securities market, so as to verify the application performance of different algorithms. In this experiment, memory index is used to reflect the efficiency of different algorithms. The smaller the memory occupied by the algorithm, the higher the efficiency of the algorithm. The unit of memory is MB. As is shown in Figure 2.
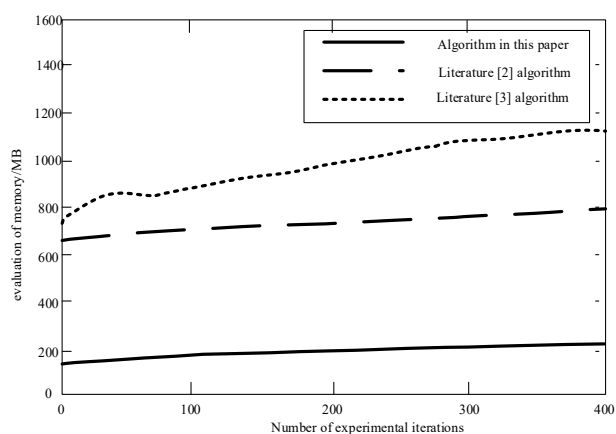

Figure 2. Comparison of evaluation of memory of different algorithms

According to the experimental data in Figure 2, during 400 experimental iterations, the fluctuation range of memory occupied by the algorithm proposed in reference [2] for big data feature investment is 600MB ~ 800MB. The fluctuation range of memory occupied by the algorithm proposed in reference [3] for big data feature investment is 600MB ~ 1200MB. Compared with the other two traditional algorithms, the algorithm proposed in this paper occupies less memory when analyzing the big data characteristics of abnormal transactions in the financial securities market. In the experimental process, the memory occupied by the proposed method is always less than 300MB. This shows that the proposed method occupies less memory in practical application and has less impact on the running efficiency of the server. Therefore, the experimental results show that the algorithm has more ideal application efficiency and better applicability.

IV. Discussion

The core of big data technology is prediction. When the scale of information data provided is large to a certain extent, the possibility of events directly reflected in the database can be predicted. With the rise of big data technology, more and more hedge funds and other institutional investors around the world have used it to determine specific investment strategies, and have also obtained more considerable investment returns than before. In this context, this paper designs a new investment analysis algorithm. The designed simulation experiments show that the average investment analysis accuracy of this algorithm is 76.91%, which is significantly higher than the literature algorithm. The big data characteristic investment analysis algorithm for abnormal transactions in the financial securities market studied in this paper can improve the accuracy of investment analysis by at least 11.24%. Compared with the other two traditional algorithms, this algorithm occupies less memory for the investment analysis of big data characteristics of abnormal transactions in the financial and securities market, and has ideal application efficiency and better applicability.

V. Conclusion

In order to improve the problems of inaccurate and inefficient analysis of investment characteristics in the traditional investment analysis algorithm, this paper improves the big data characteristic investment analysis algorithm of abnormal transactions in the financial securities market. Combined with abnormal trading and financial securities market, determine the investment strategy. The optimization objective function is established, and the genetic algorithm is used to improve the investment analysis algorithm to complete the analysis of the characteristics of abnormal trading big data in the financial securities market. Simulation experiments are designed to verify the application effectiveness of the proposed algorithm. The experimental results show that the average accuracy of investment analysis under the proposed method is improved by 11.24%, the return on investment is significantly improved, the efficiency is higher, and the situation of financial securities market is changing rapidly. In the future research, it is still necessary to deeply study the implicit rules in the trading of financial securities market in order to improve the performance of the algorithm.

References

[1] Hilliard J E, Zhang H, Webb R I. The impact of soft intervention on the Chinese financial futures market. Journal of futures markets, 2020, 40(3):374-391.

[2] Chen A, Nguyena T, Stadje M. Optimal investment under VAR-Regulation and minimum insurance. Insurance: Mathematics & Economics, 2018, 79:194-209.

[3] Tang B J, Zhou H L, Chen H, et al. Investment opportunity in China's overseas oil project: An empirical analysis based on real option approach. Energy Policy, 2017, 105(5):17-26.

[4] Tian J F, Wang D D, Kong X. Investors' attention and heterogeneity of concept stocks: influence mechanism and multidimensional measure——Matching test based on behavior big data. Jouranl of Applied Statistics and Management, 2020, 39(2):354-367.

[5] Yang J, Wen J, Jiang B, et al. Blockchain-Based sharing and tamper-proof framework of big data networking. IEEE Network, 2020, 34(4):62-67.

[6] Abawajy J H, Chowdhury M, Kelarev A. Hybrid consensus pruning of ensemble classifiers for big data malware detection. IEEE Transactions on Cloud Computing, 2020, 8(2):398-407.

[7] Sulaeman J, Wei K D. Sell-Side analysts and stock mispricing: evidence from mutual fund flow-driven trading pressure. Management Science, 2019, 15(2):782-795.

[8] Yasir M, Habib M A, Ashraf M, et al. D-GENE: Deferring the GENEration of power sets for discovering frequent itemsets in sparse big data. IEEE Access, 2020, 8:27375-27392.

[9] Chen H. Personalized recommendation system of e-commerce based on big data analysis. Journal of Interdisciplinary Mathematics, 2018, 21(5):1243-1247.

[10] Liu Z L, Hu L, Wu C Y, et al. A novel process-based association rule approach through maximal frequent itemsets for big data processing. Future Generation Computer Systems, 2018, 81:414-424.

[11] Chen Y, Guo J, Li C, et al. FaDe: A blockchain-based fair data exchange scheme for big data sharing. Future Internet, 2019, 11(11):225.

[12] Qinglai G, Bohong W, Nianfeng T, et al. Data transactions in energy internet: architecture and key technologies. Diangong Jishu Xuebao/Transactions of China Electrotechnical Society, 2020, 35(11):2285-2295.

[13] Gu Z T, Zhang X F. Research on the investment psychology difference of Chinese stock market in different trades. Journal of Guangzhou University (Natural Science Edition), 2018, 20(1):569-571.

[14] Stella V. Workflow model for big data analysis and visualization. International Journal of Applied Mathematics, Computational Science and Systems Engineering, 2021, (3): 99-106.

[15] Maddumala V, Arunkumar R. Big data-driven feature extraction and clustering based on statistical methods. Traitement du Signal, 2020, 37(3):387-394.

[16] Wang X, Zhu L, Fan Y. Transaction costs, market structure and efficient coverage of emissions trading scheme: A microlevel study from the pilots in China. Applied Energy, 2018, 220(15):657-671.

[17] Osadchyy V, Skuratovskii R V, Williams A. Analysis of the Mel scale features using classification of big data and speech signals. International Journal of Applied Mathematics, Computational Science and Systems Engineering, 2020, 2: 52-63.

[18] Chourmouziadis K, Chourmouziadou D K, Chatzoglou P D. Embedding four medium-term technical indicators to an intelligent stock trading fuzzy system for predicting: A portfolio management approach. Computational Economics, 2020(3):1152-1164.

[19] Raguseo E, Vitari C. Investments in big data analytics and firm performance: an empirical investigation of direct and mediating effects. International Journal of Production Research, 2018, 56(15-16):5206-5221.

[20] Gawankar S A, Gunasekaran A, Kamble S. A study on investments in the big data-driven supply chain, performance measures and organizational performance in Indian retail 4.0 context. International Journal of Production Research, 2019(3):1-20.