

Research on the Influence of Securities Research Newspapers on the Company's Stock Trend and Investment Strategies

Weijie Liu ^{*,#}, Hongyu Liu [#], Haojie Luo [#], Yue Zhang [#], Weiwei She [#]

College of electronics and information engineering, Guangdong Ocean University, Zhanjiang, China

* Corresponding author: a1362910274@163.com

[#]These authors contributed equally.

Abstract. Because stocks have the advantages of good liquidity and high returns, many investors will choose to invest in stocks to obtain certain income. But stocks have the disadvantage of high risk. Therefore, certain analysis and research are required when investing in stocks. Under the background of big data, research papers of securities companies can play a certain reference role for investors. The paper extracts the characteristic indexes of 10 stocks, establishes TOPSIS model and multiple linear regression model to analyze the influence of characteristic indexes on the stock trend, and proposes clear investment strategies. In the further analysis, this paper uses the gray relational analysis method to study the impact of the flash of emergencies, public opinion and natural disasters on 10 stocks, so as to propose new investment strategies. The purpose of this study is to provide investors with valuable investment strategies and improve investment efficiency.

Keywords: TOPSIS, Multiple Linear Regression, Grey Relational Analysis, Investment Strategies.

1. Introduction

The research report of the securities firm is the seller's research report (hereinafter referred to as the research report), which refers to the research report made by the researchers of the securities company on the comprehensive analysis of the value of securities and related products, or factors affecting market prices [1]. This report generally summarizes views on listed companies, industries or macro policies, and gives investment ratings on relevant stocks [2]. Therefore, research reports often contain rich and timely information, such as the business situation, major events, important information, etc. of a company, as well as the profit forecast and investment suggestions of the company, the financial forecast data, valuation results and risk tips of the company. The research report may also give relevant report ratings, historical recommendation ratings and target prices. Obviously, research papers are important reference materials for investors to make investment decisions [3]. A large amount of data also shows that investment strategies obtained based on research reports are often the most valuable. In this paper, the corresponding investment method is obtained from the information of the research report. First, the characteristic indicators are extracted from the research report. Based on these data, the deeper learning is completed, and the investment method is determined using the learning results. The factors of stock factor analysis, such as valuation factors and growth factors, are composed of these characteristic indicators. Secondly, through the use of mathematical modeling methods, and based on research reports, under the analysis of the company's financial characteristics indicators, consider the impact of external environment such as the flash of emergencies, public opinion and natural disasters on the stock trend, to better obtain the corresponding investment methods.

The contribution of this paper lies in: First, the model constructed in this paper can be applied flexibly. Secondly, the calculation of the model constructed in this paper is relatively simple, and the calculated results are also relatively reasonable. However, there are still problems that need to be improved: the stability of the multiple linear regression model still needs to be improved because the distribution of samples has a great impact on the multiple linear regression model; K-means++ algorithm is very sensitive to the setting of initial value k .

2. Research design

2.1. Model assumptions

1. This paper assumes that the data collected are accurate and effective.
2. Model assumptions of multiple linear regression:
 - (1) Suppose that the dependent variable Y is linearly related to the independent variables $X_1, X_2, \dots, X_K$;
 - (2) The variance of the residual term is the same as all the observed values: $E(\varepsilon_i^2) = \sigma\varepsilon^2$;
 - (3) The residual term is not correlated with each observation: $E(\varepsilon_i\varepsilon_j) = 0, j \neq i$.

2.2. Sample selection

This paper extracted 10 stocks, which are Changying Precision, Shunluo Electronics, Lixun Precision, Investment Accumulation, Zhongshun Jierou, Tower Group, Zhujiang Beer, Zhongju Hi tech, Radio and Television Express and Ping An. The characteristic indicators extracted from the research report include systematic risk, price earnings ratio (P/E), price to book ratio (P/B), market to sales ratio (PS), operating revenue growth rate, net profit growth rate, net sales profit rate (TTM), gross profit rate, return on net assets, total equity and total market value [4].

2.3. Symbol description

The symbols are defined in Table 1.

Table 1. Symbol definition

Symbol	meaning
d_i	alternative plan
A	decision matrix
B	normalized decision-making matrix
ω	weight
S^*_i	distance from alternative plan to positive ideal solution
VIF	variance inflation factor
R^2	goodness of fit
R^2_{adjusted}	R-squared after adjustment

3. Establishment and solution of model

3.1. TOPSIS model establishment and application

3.1.1. Introduction to TOPSIS algorithm

This paper uses TOPSIS method to calculate the corresponding score of each stock by normalizing, normalizing and standardizing the data. That is, a clear investment strategy can be proposed based on the score. The steps to use the TOPSIS algorithm are as follows:

First, the canonical decision matrix is obtained by vector programming. Let the decision matrix $A = (a_{ij})_{m \times n}$ of Multiple Attribute Decision Problems, normalized decision-making matrix $B = (b_{ij})_{m \times n}$, and:

$$b_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^m a_{ij}^2}} \quad (1)$$

$i=1, 2, \dots, m, j=1, \dots, n$. The following definitions of i and j are the same as here.

The matrix can also be changed into a normalized matrix through the standard 0-1 transformation. The optimal value of each attribute after transformation is 1, and the worst value is 0.

If x_j is an effect attribute, then:

$$b_{ij} = \frac{a_{ij} - a_j^{\min}}{a_i^{\max} - a_j^{\min}} \quad (2)$$

Next, the weighted norm matrix $C = (c_{ij})_{m \times n}$ is formed, Set the weight vector of each attribute as:

$$\omega = [\omega_1, \omega_2, \dots, \omega_n]^T \quad (3)$$

$$c_{ij} = \omega_j \times b_{ij} \quad (4)$$

Then, determine the positive ideal solution and negative ideal solution and determine the distance from each scheme to the positive and negative ideal solution. At this point, Let the j attribute value of the positive ideal solution be c_j^* , and the j attribute value of the negative ideal solution be c_j^0 . The distance from the alternative d_i to the positive ideal solution is:

$$s_i^* = \sqrt{\sum_{j=1}^n (c_{ij} - c_j^*)^2} \quad (5)$$

The distance from the alternative d_i to the negative ideal solution is:

$$s_i^0 = \sqrt{\sum_{j=1}^n (c_{ij} - c_j^0)^2} \quad (6)$$

Then, calculate the queuing index value (I. e. comprehensive evaluation index) of each scheme.

$$f_i^* = \frac{s_i^0}{s_i^0 + s_i^*} \quad (7)$$

Finally, rank the advantages and disadvantages of the schemes from the largest to the smallest according to f_i^* .

3.1.2. Data normalization

Data normalization is a basic work of data processing. Data processing is required to enhance the convenience of model establishment, improve the accuracy of the model, and increase the fault tolerance rate of the model [5]. This paper uses MATLAB software to normalize the data. The formula for data normalization is:

Table 2. Normalized matrix

0.6279	0.5667	0.5792	0.1252	1.0000	1.0000	0.7391	0.6486	0.5625	0.0076	0.0992
0.5349	0.6286	1.0000	0.8271	0.4902	0.4769	0.1739	0.0811	1.0000	0.0665	1.0000
0.6744	0.3367	0.1208	0.0273	0.4118	0.6308	0.0870	0	0.3750	1.0000	0.0518
0.1628	0.5745	0.5375	0.0455	0.4314	0	0.2609	0.7027	0.3750	0.0123	0.0810
0.6744	0	0.0375	0.0256	0.3725	0.6154	1.0000	0.7297	0.3750	0.0112	0
0	0.4969	0.2069	0.0908	0.2941	0.2154	0.6087	1.0000	0.0625	0.0209	0.0759
0.8605	1.0000	0.7319	1.0000	0	0.6462	0.3043	0.2703	0.1875	0	0.0006
0.6047	0.4873	0.2194	0.6289	0.3529	0.2308	0.6522	0.7297	0	0.0234	0.0005
0.9302	0.0314	0	0	0.3333	0.2000	0.3913	0.8108	0.3125	0.1724	0.0774

$$\frac{x - \min}{\max - \min} \quad (8)$$

See Table 2 for the matrix obtained after data normalization.

3.1.3. Positive management of data

The forward processing of data is generally to convert reverse indicators and moderate indicators into positive indicators, so it is also called the forward processing of indicators [6]. The forward processing of data can transform indicators into four types: extremely large, extremely small, intermediate and interval. This article converts all indicators into very large ones. The formula for converting very small indicators to very large indicators is:

$$\max -x \quad (9)$$

See Table 3 for the matrix obtained after data normalization.

Table 3. Positive management matrix

1.0000	0.5735	0.2181	0.0131	0.7647	0.1077	0	0.4054	0.0625	0.0113	0.0594
0.6279	0.5667	0.5792	0.1252	1.0000	1.0000	0.7391	0.6486	0.5625	0.0076	0.0992
0.5349	0.6286	1.0000	0.8271	0.4902	0.4769	0.1739	0.0811	1.0000	0.0665	1.0000
0.6744	0.3367	0.1208	0.0273	0.4118	0.6308	0.0870	0	0.3750	1.0000	0.0518
0.1628	0.5745	0.5375	0.0455	0.4314	0	0.2609	0.7027	0.3750	0.0123	0.0810
0.6744	0	0.0375	0.0256	0.3725	0.6154	1.0000	0.7297	0.3750	0.0112	0
0	0.4969	0.2069	0.0908	0.2941	0.2154	0.6087	1.0000	0.0625	0.0209	0.0759
0.8605	1.0000	0.7319	1.0000	0	0.6462	0.3043	0.2703	0.1875	0	0.0006
0.6047	0.4873	0.2194	0.6289	0.3529	0.2308	0.6522	0.7297	0	0.0234	0.0005
0.9302	0.0314	0	0	0.3333	0.2000	0.3913	0.8108	0.3125	0.1724	0.0774

3.1.4. Standardized data processing

Different evaluation indicators in this question have different dimensions and dimension units. If the data is not processed, this situation will affect the results of data analysis, thus affecting the most appropriate investment strategy. In order to eliminate the dimensional impact between indicators, this paper standardizes the original data, so that all indicators are in the same order of magnitude, making each indicator suitable for comprehensive comparative evaluation to obtain the most appropriate investment strategy.

There are 10 objects to be evaluated, and the matrix of 11 evaluation indicators is as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1\ 11} \\ x_{21} & x_{22} & \cdots & x_{2\ 11} \\ \vdots & \vdots & \ddots & \vdots \\ x_{10\ 1} & x_{10\ 2} & \cdots & x_{10\ 11} \end{bmatrix} \quad (10)$$

See Table 4 for the matrix obtained after data standardization.

Table 4. Standardized matrix using TOPSIS method

0.4669	0.3325	0.1433	0.0091	0.4705	0.0671	0	0.2058	0.0456	0.0111	0.0584
0.2931	0.3286	0.3807	0.0862	0.6153	0.6234	0.4503	0.3293	0.4104	0.0075	0.0975
0.2497	0.3645	0.6573	0.5698	0.3016	0.2973	0.1060	0.0412	0.7295	0.0653	0.9832
0.3149	0.1953	0.0794	0.0188	0.2533	0.3933	0.0530	0	0.2736	0.9827	0.0510
0.0760	0.3331	0.3533	0.0314	0.2654	0	0.1589	0.3567	0.2736	0.0121	0.0797
0.3149	0	0.0246	0.0176	0.2292	0.3837	0.6093	0.3704	0.2736	0.0110	0
0	0.2882	0.1360	0.0626	0.1810	0.1343	0.3709	0.5076	0.0456	0.0205	0.0746
0.4017	0.5799	0.4811	0.6889	0	0.4028	0.1854	0.1372	0.1368	0	0.0006
0.2823	0.2826	0.1442	0.4333	0.2172	0.1439	0.3974	0.3704	0	0.0230	0.0005
0.4343	0.0182	0	0	0.2051	0.1247	0.2384	0.4116	0.2280	0.1694	0.0761

3.1.5. Calculate score

After the data through normalization, positive management and standardization, the corresponding scores of 10 stocks can be calculated. Based on the scores, we can formulate the most appropriate investment strategy. The corresponding scores of 10 stocks are shown in Table 5 below:

Table 5. Corresponding scores of ten stocks

Stock	score
Changying Precision	3
Shunluo electron	2
Lixun Precision	4
Investment surplus	8
Zhongshun Clean&Soft	6
Tapai Group	9
Zhujiang Beer	1
Zhongju High tech	5
Radio&TV Express	10
Ping An	7

It can be seen from the table that stocks with higher scores are more suitable for investment. Since the stock of Radio and Television Express has the highest score among the 10 stocks, the most appropriate investment strategy based on TOPSIS method is to invest in one stock of Radio and Television Express.

3.2. Multiple linear regression model

This paper establishes a multiple linear regression model to analyze the impact of characteristic indicators on the stock trend. The model of multiple linear regression is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{10} x_{10} + \mu \quad (11)$$

First, test the multicollinearity regression. To test multicollinearity, we need to use the variance expansion factor (VIF) [7]. There is no serious multicollinearity in the regression model. The regression analysis table obtained by least square regression is shown in Table 6 below:

Table 6. Regression results

Evaluation quantity	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gross profit margin	14.64761	8.796055	-3.94	0.017	-59.06938	-10.22585
P/B ratio PB	6.700301	1.895037	3.54	0.024	1.438834	11.96177
P/E ratio PE	1.03818	0.2613019	-3.97	0.016	-1.763671	-0.3126899
Total market value (million yuan)	0.02786	0.0000141	-1.97	0.121	-0.000067	0.0000114
Return on net assets	20.0647	63.15113	-3.48	0.025	-395.4003	-44.72901
_cons	9.70335	10.74818	4.62	0.01	19.86163	79.54507

From the above table, we can see that the independent variable and dependent variable are most closely related to the return on equity, and the least significant is the total market value. From the regression coefficient, we can get the relationship ranking between each indicator and the evaluation object: return on equity, gross profit rate, P/B ratio, P/E ratio, and total market value.

4. Further analysis

4.1. Grey relational analysis

Grey correlation analysis is to judge whether the relationship between variables is close according to the similarity of the geometric shape of the sequence curve. The higher the similarity, the higher

the correlation between variables [8]. In this paper, grey correlation analysis is used to establish a model and study the impact of flash events, public opinion and natural disasters on stocks. The three indicators used in this paper are public opinion, illness and natural disasters: the public opinion indicator mainly takes the ratio of the number of comments that the social network platform has a positive impact on stocks on the day to the number of all comments on stocks published on the day; The disease index mainly takes the ratio of the stock purchase value on the day of epidemic infectious disease to the total stock purchase value on the day; The index of natural disaster value mainly takes the ratio of the purchase value of stocks on the day of natural disaster to the total purchase value of all stocks on that day. See Table 7 for the corresponding values of public opinion, illness and natural disasters.

Table 7. Values corresponding to public opinion, disease condition and natural disasters

Stock/index	Social network platform (Public opinion)	Epidemic infectious diseases (state of an illness)	Earthquake (Natural disasters)
Changying Precision	0.34	0.15	0.11
Shunluo electron	0.36	0.13	0.14
Lixun Precision	0.29	0.19	0.16
Investment surplus	0.13	0.23	0.32
Zhongshun Clean&Soft	0.14	0.32	0.29
Tapai Group	0.23	0.21	0.33
Zhujiang Beer	0.06	0.31	0.15
Zhongju High tech	0.3	0.28	0.12
Radio&TV Express	0.26	0.26	0.13
Ping An	0.24	0.41	0.45

4.1.1. Positive indicator

Forward processing is generally used to convert reverse indicators and moderate indicators into positive indicators, so it is also called forward processing of indicators. The forward matrix is as follows:

$$X = \begin{bmatrix} 0.28 & 0.26 & 0.34 \\ 0.30 & 0.28 & 0.31 \\ 0.23 & 0.22 & 0.29 \\ 0.07 & 0.18 & 0.13 \\ 0.08 & 0.09 & 0.16 \\ 0.17 & 0.20 & 0.12 \\ 0 & 0.10 & 0.30 \\ 0.24 & 0.13 & 0.33 \\ 0.20 & 0.15 & 0.32 \\ 0.18 & 0 & 0 \end{bmatrix} \quad (12)$$

4.1.2. Pretreatment of indicators

For this problem, the method to preprocess the indicators is to first calculate the mean value of flash, public opinion and natural disaster factor indicators related to emergencies, and then divide each element of the indicator by its mean value.

$$Y = \begin{bmatrix} 1.6000 & 1.6149 & 1.4783 \\ 1.7143 & 1.7391 & 1.3478 \\ 1.3143 & 1.3665 & 1.2609 \\ 0.4000 & 1.1180 & 0.5652 \\ 0.4571 & 0.5590 & 0.6957 \\ 0.9714 & 1.2422 & 0.5217 \\ 0 & 0.6211 & 1.3043 \\ 1.3714 & 0.8075 & 1.4348 \\ 1.1429 & 0.9317 & 1.3913 \\ 1.0286 & 0 & 0 \end{bmatrix} \quad (13)$$

4.1.3. Constructive parent sequence

The parent sequence is a data sequence that can reflect the behavior characteristics of the system. According to the matrix after the index preprocessing, the maximum value in each row of the matrix can be taken to construct a parent sequence. The parent sequence is a fictitious sequence for solving problems better. Then according to the matrix in (9), it can be easily obtained. The parent sequence for this question is:

$$X_0 = (1.6149, 1.7391, 1.3665, 1.118, 0.6957, 1.2422, 1.3043, 1.4348, 1.3913, 1.0286)^T \quad (14)$$

4.1.4. Calculation of grey relational degree

First, calculate the grey correlation coefficient matrix:

$$Z = \begin{bmatrix} 0.9777 & 1 & 0.8268 \\ 0.9633 & 1 & 0.6250 \\ 0.9259 & 1 & 0.8607 \\ 0.4760 & 1 & 0.5412 \\ 0.7322 & 0.8268 & 1 \\ 0.7066 & 1 & 0.4751 \\ 0.3333 & 0.4884 & 1 \\ 0.9115 & 0.5097 & 1 \\ 0.7241 & 0.5866 & 1 \\ 1 & 0.3880 & 0.3880 \end{bmatrix} \quad (15)$$

The grey correlation degree is the mean value of the grey correlation coefficient, so it is easy to know. The formula of the grey correlation degree (r) is as follows:

$$y(x_i) = \frac{1}{n} \sum_{k=1}^n y(x_i(k)) \quad (16)$$

After calculation, the grey correlation degree of public opinion, disease condition and natural disaster $isr_1=y(x_1)=0.7751$, $r_2=y(x_2)=0.7799$, $r_3=y(x_3)=0.7717$. The greater the grey correlation degree in this paper, the greater the impact of this indicator on the stock. From the gray correlation degree obtained, it can be seen that the disease index has the greatest impact on the stock. The weight of public opinion $w_1=0.335791177$, the weight of illness $w_2=0.325033268$, and the weight of natural disasters $w_3=0.339175555$. The score formula of the kth evaluation object is as follows: (k=1,2,..., n)

$$S_k = \sum_{i=1}^m Z_{ki} \cdot w_i \quad (17)$$

The formula for normalizing scores is:

$$S_k' = \frac{S_k}{(S_1 + S_2 + \dots + S_n)} \quad (18)$$

According to the formula, the normalized scores of 10 stocks are calculated as shown in Table 8 below:

Table 8. Normalized score

stock	Normalized score
Changying Precision	0.21355045
Shunluo electron	0.16649677
Lixun Precision	0.153850993
Investment surplus	0.142088235
Zhongshun Clean&Soft	0.124608369
Tapai Group	0.068537958
Zhujiang Beer	0.060171971
Zhongju High tech	0.050609609
Radio&TV Express	0.010835403
Ping An	0.009250242

4.2. investment strategy

K-means algorithm is a widely used clustering algorithm. Where K represents the number of class clusters; Means refers to the mean value of data objects in a cluster. It takes the mean value of all data samples in each cluster subset as the representative point of the cluster, divides the data into different categories through an iterative process, and makes the criterion function for evaluating clustering performance optimal [9]. The advantage of K-means algorithm is that the algorithm is simple and fast. When processing large data sets, the speed is high. But it also has some shortcomings. That is, the algorithm requires the user to give the number of clusters to be generated in advance, and is very sensitive to initial values and outlier data [10].

In this paper, we use the cluster model and comprehensive mathematical modeling to analyze the impact of research papers and external environmental factors on the stock trend, and propose new investment strategies. The data is processed by spss software, and standardization is carried out first. Then Table 9 and Table 10 are obtained:

Table 9. Cluster members

N	stock	cluster	distance
1	Changying Precision	1	3.152
2	Shunluo electron	3	3.231
3	Lixun Precision	3	3.018
4	Investment surplus	1	2.904
5	Zhongshun Clean&Soft	2	2.233
6	Tapai Group	2	2.867
7	Zhujiang Beer	2	2.047
8	Zhongju High tech	3	3.139
9	Radio&TV Express	2	2.004
10	Ping An	1	2.985

Table 10. Number of cases in each cluster

cluster	1	3.000
	2	4.000
	3	3.000
effective		10.000
missing		2.000

It can be seen from the chart that the 10 stocks can be divided into three categories, namely first tier stocks, which include 3 stocks: Changying Precision, Investment Surplus, and Ping An; Second tier stocks, including 4 stocks: Zhongshun Jierou, Tower Group, Zhujiang Beer, and Radio&TV Express; Third tier stocks: it includes three stocks: Shunluo Electronics, Lixun Precision, and Zhongju Hi Tech. The clustering table is shown in Table 11 below.

Table 11. Cluster

	1	2	3
Zscore (Systematic risk)	-.21284	1.91552	.22751
Zscore: (P/E)	-.45442	.09362	.54393
Zscore (P/B)	-.73985	-.47907	1.92259
Zscore (PS)	-.65256	-.48739	1.42623
Zscore (revenue growth)	-.11482	-.55170	.17060
Zscore (Net profit growth rate)	.68250	-.61583	.20349
Zscore (Net profit margin on sales (TTM))	-.97723	.64124	-.72298
Zscore (Gross profit margin)	-1.62172	1.39652	-1.39176
Zscore (ROE)	.11855	-.92933	2.25664
Zscore(Total share capital (million))	2.80638	-.36137	-.21385
Zscore(Total market value (million yuan))	-.30633	-.22675	2.82478
Zscore ((Social network platform)) Impact of public opinion)	-1.07886	-1.79809	.56511
Zscore ((Epidemic infectious disease) Effect of illness)	-.22342	.71729	-.69377
Zscore ((earthquake)Impact of natural disasters)	.84853	-.59397	-.50912

The new investment strategy is to select three stocks of the first-tier stocks: Changying Precision, Investment Promotion Surplus, and Ping A China. Among them, by looking at data column 1, we can see that the investment promotion surplus is the most suitable stock for investment among the three stocks of the first-tier stocks. The second is the four stocks of the second-tier stocks: Zhongshun Jierou, Tower Group, Zhujiang Beer, Radio and TV Express, and the last is the three stocks of the third-tier stocks: Shunluo Electronics, Lixun Precision, and Zhongmoment Hi Tech.

5. Conclusions

This paper uses the random sampling method to select and extract 10 stocks, which are Changying Precision, Shunluo Electronics, Lixun Precision, Merchants Surplus, Zhongshun Jierou, Tapai Group, Zhujiang Beer, Zhongju Hi Tech, Radio&TV Express and Ping An. The characteristic indicators extracted from the research report are respectively systematic risk, price earnings ratio (P/E), price to book ratio (P/B), market to sales ratio (PS), operating revenue growth rate, net profit growth rate, net sales profit rate (TTM), gross profit rate, return on net assets, total equity and total market value Secondly, this paper establishes TOPSIS model and multiple linear regression model, analyzes the impact of the selected 10 stocks' characteristic indicators on the stock trend, and proposes a clear investment strategy. When using the TOPSIS method, the corresponding score of each stock can be calculated by normalizing, normalizing and standardizing the data, that is, a clear investment strategy can be proposed based on the score. After calculation and analysis, the most appropriate investment strategy based on TOPSIS method is to invest in the stock of Radio TV Express. When using the multiple linear regression model, combine the characteristic indicators of the research report to analyze the characteristic indicators, and then analyze how the extracted characteristic indicators affect the stock trend. According to the multiple linear regression, the ranking of the relationship between each indicator and the evaluation object is: return on equity, gross profit rate, P/B ratio, P/E ratio, and total market value. In the further analysis, this paper uses the gray correlation analysis method to study the impact of flash events, public opinion and natural disaster factors on the selection of 10 stocks. The three indicators used in this paper are public opinion, illness and natural disasters: the public opinion indicator mainly takes the ratio of the number of comments that the social network platform has a positive impact on stocks on the day to the number of all comments on stocks published

on the day; The disease index mainly takes the ratio of the stock purchase value on the day of epidemic infectious disease to the total stock purchase value on the day; The index of natural disaster value mainly takes the ratio of the purchase value of stocks on the day of natural disaster to the total purchase value of all stocks on that day. Through the grey correlation analysis, this paper calculates the correlation degree of indicators, determines the weight of different indicators with the help of the correlation degree of indicators, and finally carries out the weighted sum score to understand the impact of various factors on the stock of securities companies. The calculated result shows that the disease index has the greatest impact on the stock. Finally, based on the idea of cluster analysis, according to the influence of the research papers of securities companies and the external environmental factors on the stock trend, this paper modifies the investment strategy, uses the spss software to cluster the stocks with K-means clustering based on the consumption amount and consumption times as the characteristic data, and divides them into first tier stocks, second tier stocks, and third tier stocks. Then it comes to the conclusion that the new investment strategy is the first choice of three stocks of the first-tier stocks: Changying Precision, Investment Surplus, and Ping An. Among them, the stock investment surplus is the most suitable stock for investment among the three first tier stocks.

References

- [1] Cao Haitao Chinese investment [M] Electronic Industry Press: 201806.275.
- [2] Zhang Shujun How to deal with a large number of investors in research papers? [N]. Chongqing Business Daily, 2017 - 12 - 18 (A04).
- [3] Lei Le, Li Ziyi, Yin Lianqian Quantitative stock selection strategy based on entropy weight method and CART decision tree [J] Special Economic Zone, 2022, (5): 114 - 118.
- [4] Yao Xiaoyu Research on the Impact of Management Power of High-tech Enterprises on the Quality of Accounting Earnings [D]. Jiangsu: Nanjing University of Information Engineering, 2021.
- [5] Liu Mengting Research on Oracle Character Recognition Based on Depth Convolution Neural Network [D] Henan: Zhengzhou University, 2020.
- [6] Zhang Yanfang Research on the financial evaluation index system of China's enterprise development ability [D] Hunan: Xiangtan University, 2010.
- [7] Jia Ling Research on the influencing factors of the demand for shared bicycles in Zhangye City based on the Logistic model [J] China Business Review, 2020, (12): 18 - 20.
- [8] Deng Yanguo, Wang Bing, Cao Zhijie, etc Spatial load forecasting of distribution network based on entropy weight method and GRA-ELM [J]. Power Engineering Technology, 2021, 40 (4): 136 - 141.
- [9] Luo Pingjuan, Zhang Shengli Logistics distribution scheme design based on K-means clustering algorithm [J] Modern Computer, 2021, 27 (24): 26 - 29.
- [10] Chen Songjing Research on intervention indicators and effect analysis methods for high-risk population of type 2 diabetes [D]. Beijing: Beijing University of Technology.
- [11] Rong, S. 2022. Proposition and modeling analysis of two asset trading strategies of gold and bitcoin under market trading. BCP Business & Management. 22, (Jul. 2022), 225 - 231. DOI: <https://doi.org/10.54691/bcpbm.v22i.1233>.