# Research on the Identification of Whitewashing Degree of Financial Statements Based on Support Vector Machine Model

Tingyu Luo

Corresponding author: m13604042455@163.com

School of management, Minzu University of China, BeiJing, China, 100081

**Abstract.** Using the data of China capital market, select the financial statement data of A-share and B-share listed companies released from 2000 to 2021 as the overall data sample, and further screen out the 5% samples with the highest and lowest degree of financial whitewashing, a total of More than 8,000 pieces of data are used as data samples for research. Then first use the improved Excess-MAD algorithm based on Benford's law to establish the identification index of financial statement whitewashing, and then select 11 financial indicators from four theoretical analysis dimensions for kendall correlation analysis, and finally retain seven modeling indicators and establish the final SVM prediction model. The results of the model run show that the support vector machine model established through this process can well identify the company samples with a relatively higher degree of financial whitewashing.

**Keywords:** Degree of whitewashing of financial statements, Benford's law, SVM model, Excess-MAD method.

## 1 INTRODUCTION

In recent years, with the exposure of individual well-known listed companies' financial fraud(Xiong,2022)[1], the identification and early warning of corporate financial fraud has also gained more attention in the academic circle. However, the research on financial statement whitewashing, which is closely related to financial fraud, is still tepid. Although compared with the financial fraud that are expressly prohibited by the laws of various countries, the negative impact caused by the whitewashing and adjustment of the public financial statements disclosed by the company's financial personnel is relatively minor, but such artificial manipulation of financial information It will also cause distortion and distortion of the financial information disclosed by the enterprise, reduce the data quality of financial statements, and then affect the use value of financial information and the judgment made by financial information users on this basis(Chen,1996)[2]. Therefore, the identification of the whitewashing degree of corporate financial statements also has certain research value.

Compared with the problem of financial fraud ,the research on the degree of whitewashing of financial statements has unique difficulties: compared with the relatively clear judgment standards for financial fraud, so far, the academic circles have still not identified the degree of whitewashing of financial statements. Lack of clear and convincing criteria for judging. However, considering that the problem of whitewashing and adjustment of corporate financial

ICBBEM 2023, May 19-21, Hangzhou, People's Republic of China
Copyright © 2023 EAI
DOI 10.4108/eai.19-5-2023.2334389

statements is very similar to the theoretical analysis process of corporate financial statement fraud, when studying the identification of the degree of whitewashing of corporate financial statements, it should be possible to start from the same perspective that is the research perspective of the relationship between financial indicators and the numerical laws of corporate financial statements. In addition, since the disclosure system of various data in my country's capital market after 2000 has gradually improved and is easy to obtain, this research will be based on the data at this stage.

This paper is mainly divided into the following parts: First, establish a statistical model using the improved Benford's law as the core, that is, the Excess-MAD algorithm(Barney,2016)[3]to analyze the financial statements of enterprises, extract useful information from a mathematical point of view, and establish a financial statement whitewash degree identification indicators. Second, using the relevant theoretical knowledge of the financial management discipline, establish a set of quantitative financial analysis indicators related to the degree of corporate financial whitewashing, and use the kendall correlation coefficient test to screen. Third, use machine learning technology to mine data, use the indicators established in the aforementioned process to build an SVM analysis model, evaluate and analyze the prediction effect of the model, and use relevant subject knowledge to explain the operation logic of the model. Finally, a machine learning evaluation model with certain practical value is obtained..

# 2 SORTING OUT RELATED RESEARCH WORK

## 2.1 Benford's law and the identification of degree of whitewashing in corporate financial statements

If we want to study the whitewashing of financial statements of enterprises, we first need to establish a reliable index to evaluate the degree of whitewashing of financial statements. Logically, since the financial statement data of an enterprise itself is an application of the carry counting system, such data must also conform to a series of inherent laws contained in the carry counting system itself. The work of Ma Boqiang et al. (Cong,2019)[4] finally proved that Benford's law is an inherent property of the carry counting system, and all data sets obtained by applying the carry counting system should conform to the description of Benford's law. In practice, Varian (Varian,1972)[5] took the lead in proposing the idea of using Benford's law to verify the practicality and reliability of data in the field of social science. In 1988, Carslaw applied Benford's law to the data quality assessment in the accounting field for the first time, which proved the feasibility of this idea(Carslaw,1988)[6]. In recent years, many Chinese scholars have studied the method of evaluating the quality of financial statements using Benford's law, and have achieved fruitful results. For example, Ding Guoyong and Feng Yu (Ding ,2003) [7] conducted Benford analysis on the financial data of colleges and universities, and believed that Benford's law can be applied to audit work. Wan Yufei et al. (Wan,2012)[8] used Benford's law to analyze the financial data of my country's listed companies in Guotai Junan database in 2011, and verified the effectiveness of Benford's law in finding signs of corporate fraud in my country's capital market. The above literature results prove that the research assumption of using Benford's law to construct an identification system for the degree of whitewashing of corporate financial statements is completely feasible, both in theoretical logic and practical application.

### 2.2 The feasibility of applying machine learning algorithms such as SVM model in financial fraud analysis model

The SVM algorithm was proposed by Vapni in 1996(Vapni,1996)[9]. Compared with other commonly used artificial intelligence methods, such as KNN, K-means, BP neural network, etc., it has a complete theoretical basis, excellent generalization ability, and powerful A series of advantages such as high-dimensional data processing capabilities. Many Chinese scholars have conducted in-depth research on the feasibility of using various machine learning models including SVM models in financial research. Chinese scholar Song Xinping(Song,2008)[10] selected A-share manufacturing companies in 2005 as the analysis object, constructed a research sample data set including 36 financial fraud companies and some manufacturing non-fraud companies, and then selected 23 financial indicators, respectively A financial fraud recognition model was constructed using a variety of machine learning methods including support vector machines. The results show that the recognition effects of these types of models are very good. Ran Maosheng et al. (Ran ,2009)[11] also constructed a SVM model based on DEA indicators and applied to financial early warning work, and proved that this type of model has a good working effect. Therefore, the research method of using the SVM algorithm to construct the research model in this study also has a solid theoretical foundation.

## 3 MODEL DESIGN

This study first establishes an evaluation index for the whitewashing degree of corporate financial statements generated by an improved algorithm based on Benford's law, and then uses this index to select a subset of research objects for model training and testing from the overall data sample. Then, referring to the research results of predecessors, a group of financial indicators that have a certain correlation with the whitewashing of corporate financial statements in financial theory is determined. Finally, the SVM model is jointly trained by using the financial index data and the evaluation index of the whitewashing degree of the financial statement.

### 3.1 Evaluation index of financial whitewashing based on Benford's law

Benford's Law, also known as the First Number Law, is an empirical description given by American physicist Frank Benford from the perspective of probability and statistics based on the discovery of American mathematician and astronomer Simon Newcomb (Benford,1938)[12]. It refers to the distribution frequency of the first digit of the data in a large number of natural distribution data sets that meet certain prerequisites and are not adjusted by human manipulation. The distribution frequency of the first digit is not evenly distributed, but a monotonous downward trend that conforms to certain mathematical laws. The mathematical formula for describing Benford's law in the decimal numbering system first derived by it is as follows.

$$P(d) = \log_{10}(1 + \frac{1}{d}) \tag{1}$$

The value range of d is a natural number, and P(d) is the occurrence probability corresponding to the number d.

The work of Ma Boqiang et al. (Cong,2019) [4] finally proved that Benford's law is not only a reasonable induction of specific mathematical phenomena, but also an objective law inherent in the carry counting system. Accordingly, it can be considered that the application of Benford's law in financial analysis has a solid enough theoretical basis.

Although Benford's law has been proven to have a high role in identifying corporate financial data whitewashing, the classic Benford's law also has the disadvantage of requiring a large number of data samples. Generally speaking, if you want to apply Benford's law for reliable data analysis, the required non-zero data samples must be greater than 5000, however, the non-zero data volume of a single enterprise report is basically about 100, which means that the classic Benford The law will make it difficult to conduct reliable in-depth analysis of data samples at the level of individual companies. In order to solve the problem that the classic Benford law cannot properly handle small sample size data, Barney. Bradley J and Schulzke. Kurt S (Barney,2016) [3] proposed an improved Excess-MAD value evaluation method based on the classic Benford law, which can well solve Small sample data analysis problems that cannot be solved by the classic Benford formula. The mathematical formula for this method is as follows:

$$\text{ExcessMAD} \approx \text{MAD} - \text{E(MAD)} \qquad (2)$$

First, $\mathrm{E(MAD)} = \sum_{k=10}^{99}\sum_{j=0}^{N}\binom{N}{j}(p_k)^j(1-p_k)^{N-j}\frac{\left|\left(\frac{j}{N}\right)-p_k\right|}{90}$ and $p_k = \log_{10}(1+\frac{1}{k})$ ;N is the sample size, that is the total number of entries of non-zero data; k is the first two digits of the data, and the value range is 10-99. In particular, when N is less than or equal to 5000, $E(MAD) \approx \frac{1}{\sqrt{158.8N}}$ .Therefore, when using this method in a data sample with less than 5000 records, you can directly use the following formula for calculation:

$$\text{ExcessMAD} \approx \text{MAD} - \frac{1}{\sqrt{158.8\text{N}}} \qquad (3)$$

And $\mathrm{MAD} = (\sum_{k=10}^{99}\frac{|\mathrm{Obs_k}-\mathrm{Exp_k}|}{\mathrm{N}})/90$ .$Obs_k$ is the actual frequency of the first two digits in the sample obtained from actual statistics, and $Exp_k$ is the theoretical probability of the first two digits calculated according to Benford's law. The interpretation method of the obtained results is also relatively simple. If the obtained value of Excess-MAD is less than or equal to 0, it means that the possibility of the data sample being manipulated is low. When it is greater than 0, it indicates that the possibility and degree of manipulation modification of the data sample are greater, and the greater the value of Excess-MAD, the greater the possibility and severity of such risks.

The work of Wang Jiamin et al. (Wang ,2022) [13] further pointed out that when using this method, the data sample size can be reduced to a collection of financial statements for a single reporting period of a single company, such as about 100 pieces of a single annual report or quarterly report. Non-zero sample size range.

## 3.2 Screening of financial indicator variable groups

According to the existing financial management theory, there are many factors that affect the company's financial situation. Therefore, when evaluating a company's financial status, decisions should not be made solely based on individual indicators, but should be comprehensively analyzed using a set of indicators with certain stability that include multiple analysis dimensions. Generally speaking, the financial indicators of enterprises can be divided into five categories: solvency indicators, operating ability indicators, growth ability indicators, profitability indicators, and cash flow indicators. By sorting out previous research results, this paper preliminarily screens out 11 financial indicators in the following four dimensions: profitability indicators: (1) return on net assets(ROE) ;(2). return on total assets(ROA) ; (3). operating profit margin ;

Indicators of solvency: (4) Current ratio ;(5) Quick ratio ;(6) Interest coverage ratio;

Operating capacity indicator : (7) Inventory turnover ratio ;(8) Accounts receivable turnover ratio ;(9) Fixed Asset turnover ratio ;(10) Total asset turnover ratio ;

Cash flow indicators: (11) Cash flow liability coverage ratio.

Then use Python to call the corr function of the pandas module for these 11 indicators, use the Kendall method to calculate the nonlinear correlation coefficient, and draw the correlation coefficient heat map shown in Figure 1, so as to eliminate some indicators with high correlation, and finally keep the table 1 The seven mold entry indicators listed.

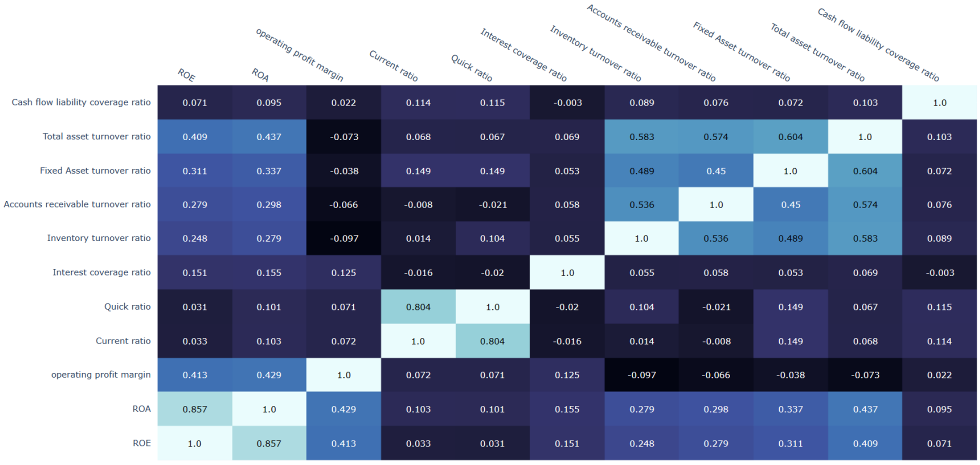

**Figure 1:** Kendall Correlation Coefficient of Original Financial Indicators: Heat Map.

**Table 1:** Definition and calculation formula of financial indicator variable group.

| Indicator dimension | Indicator name | Indicator calculation formula |
|---|---|---|
| Profitability indicator | Return on Equity (ROE) | Net Profit/Net Assets |

| | | |
|---|---|---|
| Profitability indicator | Margin Operating | Profit/Operating Income |
| Solvency indicator | Current Ratio | Current Assets/Current Liabilities |
| Solvency indicator | Interest Coverage | Earnings Before Interest and Tax (EBIT)/Interest Expense |
| Operating capacity indicator | Accounts receivable turnover ratio | Operating income/net accounts receivable |
| Operating capacity indicator | Fixed asset turnover rate | Operating income/net value of fixed assets |
| Cash flow indicator | Cash Flow Liability Coverage Ratio | Net cash flow from operating activities/current liabilities |

In addition, the seven indicators of the above four dimensions have sufficient theoretical support: The profitability indicator reflects the company's ability to obtain profits for investors, is the core indicator for evaluating the company's value, and is also the most prone to financial whitewashing operations. .The solvency indicator reflects the company's ability to pay off existing debts and obtain financing. In order to strengthen the confidence of shareholders, obtain new loans, and issue additional bonds and stocks, some companies with poor solvency indicators may have the motivation to whitewash their financial statements. The operating capacity indicator reflects the company's operating efficiency and has a certain relationship with the company's solvency and profitability. The cash flow indicator reflects the company's cash health, and it is also the most difficult type of indicator to whitewash. When the company whitewashes the relevant subjects of the profitability index, it is difficult to legally whitewash the relevant subjects of the cash flow. Therefore, it is possible to add this index to the model to assist in analyzing whether there is whitewashing in other indicators

### 3.3 Data sample selection and preprocessing process

This paper reasonably assumes that the whitewashing behavior of the company's financial statements in the current year is only directly related to the financial performance of the year. Therefore, in this paper, each company's annual report can form a piece of research data, which consists of the three statements publicly disclosed by the company in the current year (assets and liabilities Statement, cash flow statement and income statement) and the financial indicators and financial whitening indicators calculated based on the data of the three statements. The data source of this article is the raw data of all listed companies’ financial statements including A shares and B shares, parent company statements and consolidated statements in the time range of 2020-2021 available in the Guotaian database (CSMAR), a total of 101,728 strip. Afterwards, the necessary cleaning work and indicator calculations are performed on the data, and the data containing non-zero null values/infinity values in the calculated financial indicators are eliminated, and the total amount of remaining data after processing is 88,840.

Then, the normality test is carried out on the data of the whitewashing index of the financial statement. The test function is the anderson function of the scipy.stats module, and the calculated statistical value is about 1438, which is greater than the judgment value of 1.092 at the 1% level of the sample size, proving that the data It is a very typical normal distribution data. Therefore, the 5% data with the highest degree of whitewashing and the 5% data with the lowest degree of whitewashing in this sample can be directly obtained, and a total of 8983 pieces of data can be used as data samples for further training and testing. In order to make full use of the data, this study did not simply divide the training set and the test set, but used the cross-validation algorithm when training and testing the model.

In addition, in order to improve the efficiency and accuracy of model training, before officially starting to train the model, the StandardScaler function of Python's sklearn. preprocessing module is also used for data normalization, and finally the data is normalized to a mean of zero and a variance of 1 in the form of distribution. The calculation formula is.

$$X_{\text{scale}} = \frac{X - X_{\text{mean}}}{S} \qquad (4)$$

Among them, X is the original data, $X_{\text{mean}}$ is the mean value of the data, S is the standard deviation of the data, and $X_{\text{scale}}$ is the normalized data output result.

# 4 MODEL CONSTRUCTION AND RELATED PERFORMANCE ANALYSIS

## 4.1 Raw data descriptive statistics

Firstly, the descriptive statistics shown in Table 2 are performed on the original data of the two groups of data, and the comparison radar chart shown in Figure 2 is drawn. It can be found that the average value of the return on equity, operating profit margin, current ratio and interest coverage ratio of a group of data with a low degree of financial decoration is inferior to that of a data group with a high degree of financial decoration, but the turnover of accounts receivable ,the mean data of ratio, fixed asset turnover ratio, and cash flow-to-liability ratio are much better than those with a higher degree of whitewashing. And observing the variance, it can be found that the variance of each index in the data group with a lower degree of whitewashing is also much lower than that of the group with a higher degree of whitewashing, which indicates that the mold entry indicators of the group with a lower degree of whitewashing have similar rules, while the data set with a lower degree of whitewashing .The higher group has a larger difference between the entry indexes. This result is basically in line with the prediction of the relevant analysis theoretical framework of the financial management discipline, which also shows that using these two sets of data to train the SVM model may have a better judgment effect.

**Table 2:** Descriptive statistics and comparison of two groups of data

| | Roe | operating margin | current ratio | Interest coverage ratio | Accounts Receivable Turnover Ratio | Fixed asset turnover | cash flow debt ratio |
|---|---|---|---|---|---|---|---|
| Mean_1 | 0.2060 | -0.2051 | 1.5984 | -1.8192 | 259.2416 | 84.1968 | 2.4122 |
| Mean_2 | 0.2578 | 28.5561 | 3.2103 | -217.177 | 6414.193 | 6099.979 | 1.6412 |
| Std_1 | 1.2797 | 305.6216 | 1.2229 | 325.272 | 4797.287 | 2636.84 | 37.6501 |
| Std_2 | 4.5004 | 1908.74 | 8.8977 | 19935.36 | 194577 | 204038 | 52.6788 |

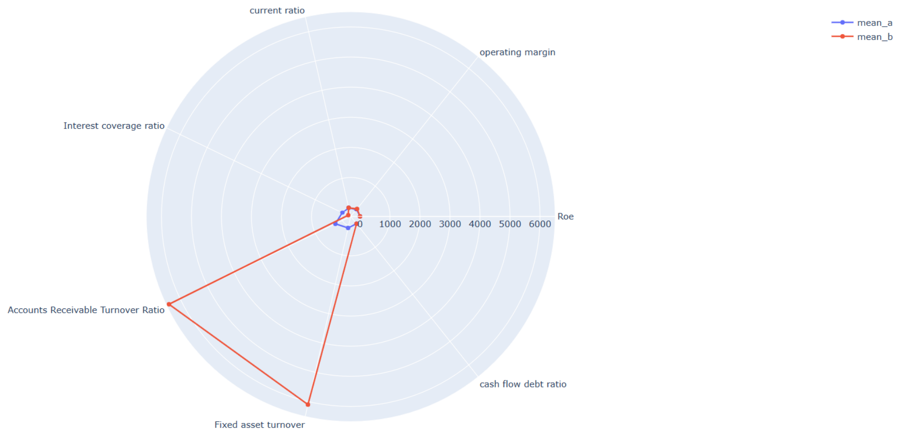

**Figure 2:** Radar chart of the comparison of the mean values of various indicators of the two sets of data.

## 4.2 SVM support vector machine model

In order to verify the effectiveness of the financial whitewashing degree SVM identification model constructed by using the financial indicators after data normalization, the model training process of this paper is as follows:

First call the function to build a linear SVM model framework, then perform K-order division on the determined 8840 pieces of data, and perform normalization processing and cross-model training successively to make full use of the data, and then use the trained model and the original data from the 8840 pieces Re-randomly screen and complete a new turn of normalized 10% data for testing, and finally obtain several evaluation indicators to determine the performance of the model, so as to obtain the financial statement whitewashing degree determination model based on financial indicators and the performance data of the model.

The process is as follows: First, use the sklearn.svm package of the Python to call the linear SVM model generator LinearSVC, and set the penalty intensity parameter to 1.0. Then use the KFold function of the sklearn.model_selection package to create a K-order cross-validation object with K = 10, and then call the cross_val_score function of the sklearn.model_selection package to run the 10-order cross-validation training process, that is, the data is randomly divided into 10 parts, each time Nine of them are used in turn as training data, and the remaining one is used as test data for model performance testing. The evaluation indicator at this stage is model accuracy. The training results of this model are shown in Table 3. It can be seen that the average accuracy of ten training sessions is about 91.20%.

**Table 3:** Model identification accuracy rate of K-order cross-validation process with K = 10.

| | Turn_1 | Turn_2 | Turn_3 | Turn_4 | Turn_5 | Turn_6 |
|---|---|---|---|---|---|---|
| accuracy | 0.9188 | 0.9155 | 0.9177 | 0.9198 | 0.8808 | 0.9198 |
| | Turn_7 | Turn_8 | Turn_9 | Turn_10 | | Mean |
| accuracy | 0.9120 | 0.9143 | 0.9020 | 0.9198 | | 0.9120 |

Then use the fit method that comes with the called svc function, use the trained model for further testing, and use the train_test_split method of the sklearn.model_selection package to randomly screen 10% of the data from the 8840 raw data used and complete the normalization After the test, the random number seed is 1 (the value of the random_state parameter is set to 1), and the output results of further tests are shown in Table 4.

**Table 4:** The results of various evaluation indicators output by the model in further testing.

| accuracy | recall | precision | F1_score |
|---|---|---|---|
| 0.99889 | 1.0 | 0.99782 | 0.99891 |

It can be seen that when the trained model is used to predict real data, very good prediction results can be obtained, and the operating logic of the model is roughly in line with the theoretical framework of the discipline based on the analysis of corporate financial statements, so the model has relatively good interpretability. In summary, the SVM judgment model obtained in this study based on financial indicators training can better judge the degree of whitewashing of corporate financial statements at a certain level of accuracy and understandability, and has certain application value. This is also the greatest significance of this article.

## 5 CONCLUSION

This paper takes listed companies in mainland China from 2000 to 2021 as the research object, and uses the improved Benford calculation formula to initially identify the degree of whitewashing of corporate financial statements, and then selects the financial statement data accounting for 10% of the total. After theoretical screening and Kendall correlation After the coefficients are screened, the SVM model is established through normalization. The empirical results show that the trained linear SVM model has excellent recognition accuracy, and compared with using the mathematical probability model based on the improved Benford's law, directly applying the trained SVM model for report evaluation has a series of advantages such as short judgment time, less manual intervention, less input data, more in line with the analysis framework of classical financial theory, and high comprehensibility of model operation result data. And it also shows that it is not only feasible but also effective to use the SVM algorithm to build a model when studying the whitewashing degree analysis of corporate financial statements. Since the practical research in the field of intelligent financial analysis, especially in the identification of corporate financial statements is still relatively weak in my country at this stage, this study also has a certain practical value.

In the next step of research, we can further explore on the basis of improving the operating efficiency of the model, and further improve the accuracy of the model to determine the data in the middle area by expanding the capacity of the training data set and the model indicators, thereby improving the practicability of the model. In addition, it is also possible to explore algorithms from this basic model to derive more new models, such as using a nonlinear SVM model for fitting, or using the bee colony algorithm to improve the parameter selection of the model, etc. The above-mentioned technical comprehensive improvements to the SVM model are also valuable research directions for further research.

## REFERENCES

[1] Xiong Fangjun, Zhang Longping, Han Yue. Research on Financial Fraud Risk Identification and Governance Countermeasures of Listed Companies——Taking Luckin Coffee as an Example [J]. Friends of Accounting, 2022, No.675(03): 55-61. https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C44YLTlOAiTRKibYlV5Vjs7iJTKGjg9uTdeTsOI_ra5_XV4Sociuoi3KjaAX07VB1IrYqmFSZeJ_UlDcLrxVTejI&uniplatform=NZKPT

[2] Chen Hanwen, Ding Peng. Discussion on Several Issues of Financial Whitewashing [J]. Chinese Agricultural Accounting, 1996(06): 6-7 https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C44YLTlOAiTRKjkpgKvIT9NkZNmQNo4kSVptbzpA-48rhvZYbwIJcCkLiAyeuui7ymstBq1RzR8wZFJXsg6_Q_QD&uniplatform=NZKPT.

[3]Barney, Bradley J.; Schulzke, Kurt S. (2016). Moderating “Cry Wolf” Events with Excess MAD in Benford's Law Research and Practice. Journal of Forensic Accounting Research, 1(1), A66–A90. DOI:10.2308/jfar-51622.

[4] Cong M , Li C , Ma B Q . First digit law from Laplace transform[J]. Physics Letters A, 2019. DOI: 10.1016/j.physleta.2019.03.017

[5]Hal V. Benford’s law [J]. American Statistician, 1972, (26): 65. DOI:10.1080/00031305.1972.10478934.

[6]Carslaw CPAN. Anomalies in Income Numbers:Evidence of Goal Oriented Behavior [J] .The Accounting Review,1988,（2）:321-327 https://www.jstor.org/stable/248109

[7] Feng Yu, Ding Guoyong. Banford's Law and Its Auditing Application [J]. Auditing:theory & Practice,2003(12):44-45..https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C44YLTlOAiTRKgchrJ08w1e7ZCYsl4RS_3guY8q7VlFK3kJkUrYVn0z8JPkKvT-9u5OZLki_7cjjSEDgp2FE8Sps&uniplatform=NZKPT

[8] Chen Xi, Wan Yufei, Li Lu. The Applicability of Finding Corporate Fraud Based on Benford's Law——An Empirical Test on the Financial Data of my country's Listed Companies [J]. Finance and Accounting Monthly,2012,No.610(06):45-48.DOI:10.19641/j.cnki.42-1290/f.2012.06.013.

[9]Vapnik V N. The nature of statistical learning theory [M]. Berlin:Springer-Verlag,1995:1 － 50.

[10] Song Xin-ping,Ding Yong-sheng,Zhang Ge-fu. Application of integrated classification method in identifying risk of fraudulent financial report [J]. Computer Engineering and Applications, 2008, 44(34):226 - 230. https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C44YLTlOAiTRKgchrJ08w1e7VSL-HJEdEx3uP3Vz1W4QmuDdV_GeHg5RPz8at3_RM5FVxTZtsbbJ1l5BC9dyDHxK&uniplatform=NZKPT

[11]Ran Maosheng, Zhou Shu, Huang Lingyun. Application of SVM Model Based on DEA Index in Financial Early Warning [J]. Statistics & Decision, 2009, No.296(20): 143-145. https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C44YLTlOAiTRKgchrJ08w1e75TZJapvoLK1dyuQnPmdvsY6TJ_GK366pvU68mlz-env5OPHeHSfeS9eBdxKveJJg&uniplatform=NZKPT

[12]Benford F.The law of anomalous numbers [J]. Proceedings of the American Philosophical Society, 1938, 78(4): 551-572. https://www.jstor.org/stable/984802

[13]Wang Jiamin, He Ding. Financial Fraud Risk Governance Effect of Actively Shorting Chinese Concept Stocks——A Case Study Based on Benford's Law Compliance Test [J]. The Chinese Certified Public Accountant, 2022, No.280(09): 42-46.DOI:10.16292/j.cnki.issn1009-6345.2022.09.028.