Financial Economics Insights

Article

Financial Fraud Detection Study - Based on Logit Model

Yangyang Wan [1,*]

[1] Department of Finance, Anhui University of Finance & Economics, Bengbu, China
* Correspondence: Yangyang Wan, Department of Finance, Anhui University of Finance & Economics, Bengbu, China

Abstract: The problem of financial fraud has always been a key issue of concern for researchers at home and abroad, and financial fraud of listed companies occurs from time to time at home and abroad, which produces a huge loss of interest for investors as well as the downturn and instability of the capital market, for this reason, it is necessary to carry out a study on the detection of financial fraud. This paper analyzes the financial fraud detection of Chinese listed companies by establishing a Logit model, firstly, obtain the financial statement sample data of Chinese listed companies from CSMAR, and divide the sample into training set and test set with the ratio of 9:1 to preprocess the sample data and fill in the missing values; secondly, this paper selects the financial statement data, and based on the previous research, selects the characteristics that are related to the risk of financial fraud and constitute the risk of financial fraud. high features, and constitute the feature indicators for financial fraud detection; again, since detecting financial fraud is a binary classification problem, the sample is divided into fraudulent companies and normal companies, so this paper studies the financial fraud problem through the discrete choice model in the econometric model, constructs a Logit model, conducts a goodness-of-fit test on the sample data of CSMAR, and estimates the parameters using the method of maximum likelihood estimation ; finally, the test set data is used to test the model's ability to predict financial fraud.

Keywords: logit model; financial fraud; discrete choice model; binary classification

Published: 30 October 2024

Copyright: © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Background of the Study

Since the reform and opening up, China began to build the capital market, from scratch, from the Shanghai Stock Exchange and Shenzhen Stock Exchange to the multi-level capital market, and currently has been committed to the development of the multi-level capital market, to meet the financing needs of listed companies, and to promote the high-quality development of China's economy; however, domestic and foreign cases of financial fraud occur from time to time, and in vestors suffered huge economic losses, and capital market lose confidence in the capital market, and cause harm to the capital market, triggering volatility. For example, the Enron incident, World Com, and Weilang Pharmaceuticals, etc. Since the establishment of the domestic stock exchange, there have been a lot of financial frauds of listed companies, such as Kangmei Pharmaceuticals, Rising Star Coffee, and Leshi, etc. These companies are not afraid of the national laws and regulations. These companies defy national laws and make false reports on the content of financial reports to deceive investors, regulators and auditors in order to maximize their own interests. In fact, by seeking econometric models, conducting research on financial fraud detection, predicting listed companies with serious financial problems beforehand, and predicting fraudulent companies, it is possible to avoid the loss of investor's interests and the risk of volatility in the capital market. Based on this, the study of financial fraud in listed companies is of great practical importance for risk management and capital market.

Since the financial fraud problem is a binary classification problem, there are two aspects of using econometric models to study the financial fraud detection problem, first, generally traditional econometric models assume that the random error term obeys a normal distribution and satisfies the conditions of zero mean, homoskedasticity, and absence of autocorrelation, whereas in a binary classification problem, the random error term does not satisfy the conditions of normal distribution as well as homoskedasticity; second, the traditional econometric models the explanatory variables are continuous variables, in the binary classification problem there are only two outcomes of the explanatory variables, fraudulent companies and normal companies. Based on this, so much so that there is a serious bias in modeling using this type of traditional models, thus failing to predict financially fraudulent companies. For this reason, this paper investigates financial fraud based on a Logit model with discrete choice modeling and with CSMAR sample data.

1.2. Literature Review

For the financial fraud problem, many scholars have carried out a very profound study, theregulatory authorities promulgated a lot of laws and regulations to prevent corporate financial fraud, cheating investors and disrupt the market; domestic and foreign scholars are committed to rese arching the causes of financial fraud, financial characteristics related to the risk of fraud, the detection of financial fraud, and analyzing and researching the problem of financial fraud from different perspectives.

Suh et. al. (2018) explored how finance executives retroactively account for crossing the line into financial statement fraud when acting or reacting to a financialized corporate environment. Findings Abnormal cash flow from operations (CFO) and abnormal production costs are used as proxies for true earnings management. Nasir (2018) finds that financial statement fraud firms are involved in manipulating production costs in the two years prior to the fraud. Reurink et. al (2018) describe the empirical range of financial fraud documented in the academic literature. In light of recent studies identifying certain interest groups as potential whistleblowers. Smaili et. al (2019) proposed a comprehensive conceptual framework that examines whistleblower behavior by type of whistleblower. Nasir et. al (2019) aimed to examine the relationship between the presence of Malay directors in Malaysian boards of directors and financial statement fraud. Rengganis et. al (2019) aimed to examine the relationship between the presence of Malay directors in Malaysian boards of directors and financial statement fraud. al. 2019) aims to analyze the role of elements of fraud diamond in detecting financial statement fraud by examining the effect of variables on financial statement fraud. The study aimed to prove the research hypothesis that there is an effect of financial ratios consisting of profitability, leverage and liquidity on the risk of financial statement fraud and that auditor quality moderates the relationship between financial ratios and financial statement fraud. Albashrawi (2021) aims to review research on the use of data mining tools to detect financial fraud over a ten-year period and to communicate current trends to academic scholars and industry practitioners. Other influential works include (Uwuigbe et al., 2019). Guided by Systemic Functional Linguistics (SFL) theory, (Dong et. al, 2018) proposed an analytical framework to mine unstructured data from financial social media platforms to assess the risk of corporate fraud. Since older directors are typically more experienced and have more to lo se if they fail to fulfill their oversight responsibilities, (Xu et. al, 2018) expect them to be more competent and have a stronger incentive to closely monitor the CEO. Thus (Xu et. al, 2018) propose that when the average age of the board increases (i.e., board age), the CEO's involvement in corporate financial fraud is less likely. The independent variables are variable pressures represented by financial stability, external pressures, and financial goals, opportunity represented by the nature of the industry, rationalization represented by total accruals, competence represented by changes in directors, and arrogance represented by management ownership (Evana et. al.) (Raval et. Al,2019) aim to analyze the attributes of Ponzi schemes ("Ponzis") in order to determine whether they whether they are a unique class of financial fraud. The theme of (Homer, 2019) is to examine the existing literature on the fraud triangle. (Kusaya et. al, 2020) develops empirical models for detecting insider abuse and fraud occurring in U.S. commercial banks with the goal of identifying leading indicators of fraud and fraud prediction. Fraud is still occurring. Therefore, this study aims to develop a multidimensional theoretical model that explains the factors that influence leverage of control (LOC) (categorized as attitudes, subjective norms, and perceived behavioral control); how the level of leverage of control affects the fraud prevention framework, and whether leverage of control mediates the relationship between planned behavioral variables and an effective fraud prevention framework (Rosli et. al, 2020). Evidence of a strong negative association between consumer fraud victimization and individuals' perceptions of their financial situation is provided u sing U.S. household panel data (Brenner et. al, 2020). (Al-Hashedi et. al, 2021) shows a list of countries exposed to financial fraud.

Based on the above analysis, this paper studies the financial fraud detection problem by establishing Logit model. Logit model can well predict the financial fraud of listed companies, thus helping investors to understand the business situation and future development of the enterprise, an d facilitating investors to more accurately judge the risk and make rational investment behavior.

2. Model Building

2.1. Logit Model

The traditional linear model for solving the binary discrete choice problem there are errors in the regression results, the original true value of the value taken only 0 and 1 two results, if the linear model is used to get the fitted straight line, there will be less than 0 and greater than 1 results, for such a result, we cannot explain the results, and therefore cannot be used in the original linear model. For the dependent variable is a binary value, this paper adopts the logit model in the discrete choice model and uses the great likelihood estimation method to estimate the parameters of the model.

Logit model derivation process:

An auxiliary continuous variable is first identified by constructing an equation to model

the binary discrete dependent variabley_i^*:

$$y_i^* = x_i'\beta + \varepsilon_i (i = 1,2,\ldots,n) \quad (1)$$

Then build the relationship with the dependent variable y_i^*through a segmented function:

$$y_i = \begin{cases} 1 & y_i^* > 0 \\ 0 & y_i^* \le 0 \end{cases} \quad (2)$$

According to equation (1) and equation (2) then the probability can be introduced:

$$\begin{aligned} p(y_i|x) &= p(x_i'\beta + \varepsilon_i > 0) \quad (3) \\ &= p(\varepsilon_i > -x_i'\beta) \\ &= 1 - p(\varepsilon_i \le -x_i'\beta) \\ &= p(\varepsilon_i \le x_i'\beta) \end{aligned}$$

Assuming that the residuals obey a logistic distribution and satisfy the assumption of independent homogeneous distribution, when the mean of the residualsε_i is 0 and the variance is$\pi^2/3$, Equation (1) and Equation (2) constitute the general form of the Logit model:

$$p(y_i|x) = \Lambda(x_i'\beta) = exp(x_i'\beta)/(1 + exp(x_i'\beta)) \quad (4)$$

In summary, equation (4) is the equation of the binary Logit model.

2.2. Parameter Estimation of the Model Using Maximum Likelihood Approach

Let $p$the probability of $y_i = 1$, then $1-p$the probability of$y_i = 0$, the distribution law of the binary Logit model is$p^{y_i}(1-p)^{1-y_i}$, according to formula (4) then the likelihood function can be constructed as follows:

$$L(\beta;x) = \prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i} \qquad (5)$$
$$= (\frac{exp(x_i'\beta)}{1+exp(x_i'\beta)})^{y_i}(1-\frac{exp(x_i'\beta)}{1+exp(x_i'\beta)})^{1-y_i}$$

Considering the existence of exponentials in the formula on the right side of the equation, it is more difficult to solve, so taking logarithms on both sides, the log-likelihood function can be obtained.

$$LnL(\beta;x) = Ln\prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i} \qquad (6)$$
$$= \sum_{i=1}^{n} y_i \cdot ln(\frac{exp(x_i'\beta)}{1+exp(x_i'\beta)}) + \sum_{i=1}^{n}(1-y_i)\cdot ln(\frac{1}{1+exp(x_i'\beta)})$$

By solving Eq. (6) for the maximum value, the parameters to be estimated for the binary Logit model can be solved.

2.3. Validity, Goodness of Fit, and Coefficient Interpretation of Binary Logit Models

2.3.1. Validity Test of the Model

The validity of a binary Logit model can be tested by the likelihood ratio statisticΩ, which compares the relative magnitude of the likelihood function value (L_0) for a model that includes only the constant term model to the likelihood value (L) for a model that includes all explanatory variables as well as the constant term. The test formula is as follows:

$$\Omega = -2(ln\,L_0 - ln\,L) \sim x^2(K-1) \qquad (7)$$

where the statistic Ωobeys a distributionx^2, whose degree of freedom (K-1) is the difference between the number of current model parameters and the number of parameters in a model containing only constant terms.

2.3.2. Judgement of Model Fit Goodness of Fit

The most commonly used metric for judging the goodness-of-fit of Logit models is$McFadden'sR^2$,also known as pseudoR^2. The basic idea is similar to the likelihood ratio test, in that it involves comparing the relative size of the likelihood value (L_0) of a model that contains only the constant term with the likelihood value (L) of a model that contains all the explanatory variables as well as the constant term. The formula for this indicator is as follows:

$$McFadden'sR^2 = 1\text{-}\frac{ln\,L}{ln\,L_0} \qquad (8)$$

In general, a pseudo-indicatorR^2 of 0.2 or higher indicates goodness of fit, but if the focus is on analysing the influencing factors in the actual application, it is possible to focus less on this indicator.

2.3.3. Interpretation of Model Coefficients

According to the previous definition, the probability of$y = 1$ is$p$, then the probability of$y = 0$ is $1-p$, the probability ratio is$\frac{p}{1-p}$, and the expression of the binary Logit model is:

$$\frac{p}{1-p} = \frac{exp(x_i'\beta)}{1+exp(x_i'\beta)}/(1-\frac{exp(x_i'\beta)}{1+exp(x_i'\beta)}) \qquad (9)$$
$$= exp(x_i'\beta)$$

Taking the logarithm of both sides, we can get the linear form of the binary Logit model:

$$ln\frac{p}{1-p} = x_i'\beta \qquad (10)$$

When the explanatory variable is continuous, the regression coefficient $\hat{\beta}_j$ means that an increase in a small amount of the variable x_j causes a marginal change in the logarithmic probability ratio when the explanatory variable is continuous. When the explanatory variable is a discrete variable, assuming that x_j increases by one unit from x_j to x_j+1, and the new value p of the probability of the event is p^*, then the ratio of the new probability to the original probability is:

$$\frac{\frac{p^*}{1-p^*}}{\frac{p}{1-p}} = \frac{exp[\beta_1+\beta_2 x_2+\ldots+\beta_j(x_j+1)+\ldots+\beta_k x_k]}{exp[\beta_1+\beta_2 x_2+\ldots+\beta_j x_j+\ldots+\beta_k x_k]} = exp(\beta_j) \quad (11)$$

3. Empirical Analysis

3.1. Variable Settings

In the study of financial fraud detection, the financial data of listed companies are usually used for detection, according to the relevant research of previous research, this paper analyzes the financial statements of listed companies from the four aspects of solvency, development ability, profitability, and operating ability, and the financial variables in the financial statements of listed companies and financial ratios are utilized to study the detection of financial fraud, and the analysis is carried out using 17 financial indicators, respectively are current ratio, quick ratio, cash ratio, equity ratio, net cash flow from operating activities/total liabilities, accounts receivable turnover, inventory turnover, current assets turnover, fixed assets turnover, return on net assets, return on invested capital, net operating margin, operating profit before interest and tax, cash-to-total-profit ratio, total assets growth rate, return on net assets growth rate, and net profit growth rate.

3.2. Data Sources

In view of the completeness, reliability and authenticity of the data, this paper obtains the relevant data on the financial statements of listed companies from the China Economic and Financial Research Database (CSMAR), and selects the sample data from the financial annual reports from 2011 to 2021. There are two problems in the selection of data, first, because there is a certain time lag from the financial fraud to the discovery and publication of the penalty results by the CSRC and other organizations, it may lead to the possibility that the listed companies closer to the present time may be fraudulent but have not been detected. Therefore, when training the model, the sample data from 2011 to 2018 is used, and the sample data from 2018 to 2021 is used to test the prediction accuracy of the model; secondly, since the number of non-fraudulent samples in reality is much higher than the number of fraudulent companies, resulting in an imbalance in the data, which is likely to affect the prediction accuracy of the model, in order to circumvent such a problem, this paper, in the choice of the number of 74 samples of fraudulent companies and 74 samples of non-fraudulent companies are used to make the model better classify the characteristics of fraudulent companies.

3.3. Statistical Characterization of Data

In order to have a preliminary understanding of the financial data of the listed companies, first of all, the financial data are statistically described in STATA.16 software and the following are the basic characteristics of the financial data:

Statistical description can be understood: the standard deviation of the four financial indicators of accounts receivable turnover, inventory turnover, fixed asset turnover, and net profit growth rate are 403.4336, 39.01824, 97.91204, and 19.88346 respectively, which are relatively high compared to the standard deviation of the other financial indicators, and it is possible to have an impact on the parameter estimation of the results. To address this issue, the usefulness of these four indicators is determined by comparing the results of parameter estimation with or without these four indicators. (Figure 1)

Variable	Obs	Mean	Std. Dev.	Min	Max
fraud	148	.5	.5016978	0	1
f010101a	148	1.552895	1.292611	.021511	9.372211
f010201a	148	1.077037	1.235269	.010682	9.368977
f010401a	148	.432949	.4955643	.000417	2.579489
f011701a	148	1.312754	2.422525	-9.15563	15.61895
f012301b	148	.0805576	.2448022	-.88604	1.555571
f040205c	145	101.9183	403.4336	.292838	3662.259
f040505c	148	12.62249	39.01824	.012908	310.3083
f041205c	148	1.348799	1.037603	.013481	4.910961
f041405c	148	18.14958	97.91204	.202106	1127.887
f050504c	138	.0199809	1.0018	-11.25531	2.528886
f051201b	141	.0563342	.1398653	-1.029876	.582054
f051501c	148	.658856	9.100702	-10.94489	109.7486
f052401c	148	1.092313	12.47655	-6.692657	151.4249
f052901c	130	-.4204743	12.43749	-130.1541	19.48702
f080602a	148	.1572911	.5244637	-.735768	4.904927
f080702b	129	-.5790662	2.57491	-18.9074	4.555302
f081002b	128	-2.854735	19.88346	-180.2037	5.161247

Figure 1. Statistical characterization of data.

The financial data of listed companies are analyzed by the Logit model, analyzed by STATA.16 software, and the parameters of the model are estimated by the maximum likelihood estimation method, and the empirical results are as follows (Figure 2) :

```
Logistic regression                             Number of obs     =        115
                                                LR chi2(17)       =      40.20
                                                Prob > chi2       =     0.0012
Log likelihood =  -54.80769                     Pseudo R2         =     0.2684
```

fraud	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
f010101a	.281986	.63832	0.44	0.659	-.9690983	1.53307
f010201a	-.1894713	.7463724	-0.25	0.800	-1.652334	1.273392
f010401a	-.5135049	.9459243	-0.54	0.587	-2.367482	1.340473
f011701a	.2976071	.3354814	0.89	0.375	-.3599244	.9551387
f012301b	-1.626953	1.731124	-0.94	0.347	-5.019894	1.765988
f040205c	-.0015228	.0028239	-0.54	0.590	-.0070577	.004012
f040505c	-.0204846	.0287234	-0.71	0.476	-.0767814	.0358123
f041205c	-.9902364	.4799998	-2.06	0.039	-1.931019	-.049454
f041405c	-.0023778	.0195709	-0.12	0.903	-.040736	.0359804
f050504c	15.93838	10.13278	1.57	0.116	-3.921514	35.79827
f051201b	-28.21698	17.91327	-1.58	0.115	-63.32635	6.892389
f051501c	14.0436	8.215345	1.71	0.087	-2.05818	30.14538
f052401c	-11.76223	7.287656	-1.61	0.107	-26.04577	2.521316
f052901c	-.0442601	.0554088	-0.80	0.424	-.1528594	.0643392
f080602a	-1.659618	1.257924	-1.32	0.187	-4.125103	.8058667
f080702b	-.1285126	.9936425	-0.13	0.897	-2.076016	1.818991
f081002b	.0242082	.8040233	0.03	0.976	-1.551648	1.600065
_cons	1.500472	1.311855	1.14	0.253	-1.070716	4.07166

Figure 2. The result of the parameter estimation.

The second exclusion of indicators with relatively high standard deviation and parameter estimation of the remaining indicators yielded the estimated results as shown below (Figure 3) :

Logistic regression		Number of obs	=	123
		LR chi2(11)	=	45.73
		Prob > chi2	=	0.0000
Log likelihood = -59.832809		Pseudo R2	=	0.2765

fraud	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
f010101a	.1936594	.5380528	0.36	0.719	-.8609048	1.248223
f010201a	.3452778	.654392	0.53	0.598	-.9373069	1.627863
f010401a	-.8543065	.9120784	-0.94	0.349	-2.641947	.9333343
f011701a	.2964022	.2860407	1.04	0.300	-.2642273	.8570317
f012301b	-2.310342	1.349075	-1.71	0.087	-4.954481	.3337974
f041205c	-.9069545	.3609857	-2.51	0.012	-1.614473	-.1994355
f050504c	16.40164	8.492945	1.93	0.053	-.2442257	33.04751
f051201b	-31.53519	16.00507	-1.97	0.049	-62.90454	-.165838
f051501c	.9838279	1.385932	0.71	0.478	-1.732548	3.700204
f080702b	-.3319001	.2958831	-1.12	0.262	-.9118202	.2480201
f080602a	-1.337042	1.169449	-1.14	0.253	-3.62912	.9550364
_cons	.7379096	1.05287	0.70	0.483	-1.325678	2.801497

Figure 3. The parameter estimation results of the standard deviation indicator are not included.

From the size of the pseudo R^2 and the significance of the parameters of the indicators, it is known that the second estimation is slightly better than the results of the first estimation, in the analysis of the results of the parameter estimation, although some of the parameters of the indicators are not very significant, but they are also explained here. When other conditions are not tabulated, a one-unit increase in the current ratio causes a 19% increase in the log odds ratio of fraud occurring; a one-unit increase in the quick ratio causes a 34% increase in the log odds ratio of fraud occurring; a one-unit increase in the cash ratio causes an 85% decrease in the log odds ratio of fraud occurring; a one-unit increase in the equity ratio causes a 29% increase in the log odds ratio of fraud occurring; and net cash flow/total liabilities generated from operations is a one-unit increase in the log odds ratio of fraud occurring. Net cash flows from operating activities/total liabilities increased by one unit decreased by 2.3 times the logarithmic odds of fraud occurring; current asset turnover increased by one unit decreased by 90% of the logarithmic odds of fraud occurring; return on net assets increased by one unit increased by 16 times the logarithmic odds of fraud occurring; return on invested capital increased by one unit decreased by 31 times the logarithmic odds of fraud occurring; net operating profit margin An increase of one unit in the net operating margin causes a 98% increase in the log odds ratio of fraud occurrence; an increase of one unit in the growth rate of total assets causes a 33% decrease in the log odds ratio of fraud occurrence; and an increase of one unit in the growth rate of return on net assets causes a 1.3 times decrease in the log odds ratio of fraud occurrence. The results of the analysis also show that the parameter significance of the four indicators of net cash flow from operating activities/total liabilities, return on net assets, current asset turnover, and return on invested capital are significant at the 10% level.

3.4. Validity of Logit's Model Goodness of Fit

The values of the likelihood ratio statistics LR chi2(17) and LR chi2(11) are 40.2 and 45.73, respectively, corresponding to p-values much less than 0.01, indicating that the joint significance of all the coefficients of the whole equation (except for the constant term) is very high, which suggests that the model is a valid one.

The pseudo R^2 of the two parameter estimates are 0.2684 and 0.2765, respectively, and it can be seen that the pseudo R^2 values of the two times are greater than 0.2, which indicates that the model has a better goodness of fit.

3.5. Evaluation of Predictive Accuracy of Empirical Results of Logit

Models (Figure 4):

Logistic model for fraud

Classified	True D	True ~D	Total
+	22	11	33
-	19	63	82
Total	41	74	115

Classified + if predicted Pr(D) >= .5
True D defined as fraud != 0

Sensitivity	Pr(+| D)	53.66%
Specificity	Pr(-|~D)	85.14%
Positive predictive value	Pr(D| +)	66.67%
Negative predictive value	Pr(~D| -)	76.83%
False + rate for true ~D	Pr(+|~D)	14.86%
False - rate for true D	Pr(-| D)	46.34%
False + rate for classified +	Pr(~D| +)	33.33%
False - rate for classified -	Pr(D| -)	23.17%
Correctly classified		73.91%

Figure 4. Evaluation of the accuracy of model predictions**.**

From the results of the predictive analysis, it can be seen that here the critical point of prediction classification when 0.5, the probability of classification prediction is greater than 0.5, the company is considered to be fraudulent, when the probability of prediction is less than 0.5, it is considered that the company is not fraudulent. For the Logit model, the prediction accuracy for non-fraudulent companies is 66.67% and for fraudulent companies is 76.83%, with a total sample accuracy of 73.91%.

4. Conclusions and Recommendations

4.1. Conclusion

This paper conducts empirical analysis and research on financial fraud of listed companies in China, comprehensively selects 17 representative financial indicators, analyzes the detection of financial fraud of listed companies in China by constructing a binary classification model, and makes suggestions on how to reduce the behavior of financial fraud of listed companies in China. In this paper, the financial fraud penalty announcements of listed companies publicly released by the China Securities Regulatory Commission between 2011 and 2021 are used as the sample data for the study of this paper, and at the same time, in order to avoid the interference of unbalanced data on the prediction accuracy, this paper adopts the positive and negative samples1:1 , and the number of positive samples is 74, and the number of negative samples is 74. After that, model regression is performed on software STATA.16 for financial indicators, and finally out-of-time samples are selected for testing. The following conclusions are drawn:

According to the results of the empirical analysis, it can be seen that the four indicators of net cash flow from operating activities/total liabilities, return on net assets, current asset turnover and return on invested capital have a significant role in analyzing the financial fraud of listed companies, and are of greater help in detecting the financial fraud of listed companies, and the auditing institutions and regulatory authorities can focus on analyzing these four financial indicators of listed companies when reviewing their financial statements. Auditors and regulators can focus on analyzing these four financial indicators of listed companies when reviewing the financial statements of listed companies.

Logit model is a binary classification model, for financial fraud detection, has a better classification effect, this paper uses Logit model for the empirical analysis of financial fraud detection of listed companies, the classification prediction accuracy is as high as 73.19%, for the listed company's financial fraud has a better classification for the regulatory authorities and auditing institutions to review the listed company's financial statements to provide new help.

4.2. Recommendations

First, auditing departments and regulators can use the Tobit model when reviewing the financial statements of listed companies to study and analyze whether the listed companies are fraudulent.

Secondly, when reviewing the financial statements of listed companies, the auditing and regulatory authorities can focus on the four indicators of net cash flow from operating activities/total liabilities, return on net assets, turnover of current assets, and return on invested capital, which can be of great help in detecting the detection of financial fraud in listed companies.

Third, investors can draw on the above analysis to analyze the financial statements of listed companies to avoid unnecessary risks.

References

1. Dwi Ratmono; Yuvita Avrie Diany; Agus Purwanto; "Dapatkah Teori Fraud Triangle Menj elaskan Kecurangan Dalam Laporan Keuangan?", Jurnal Akuntansi Dan Auditing, 2018.
2. Ikseon Suh; John T. Sweeney; Kristina Linke; Joseph M. Wall; "Boiling The Frog Slowly: The Immersion Of C-Suite Financial Executives Into Fraud," Journal Of Business Ethics, 2 018.
3. Noorul Azwin Binti Md Nasir; Muhammad Jahangir Ali; Rushdi M.R. Razzaque; Kamran Ahmed; "Real Earnings Management And Financial Statement Fraud. Evidence From Mala ysia", International Journal Of Accounting & Information Management, 2018.
4. Arjan Reurink; "Financial Fraud: a Literature Review", Consumer Financial Fraud Ejourna l, 2018.
5. Nadia Smaili; Paulina Arroyo; "Categorization Of Whistleblowers Using The Whistleblowi ng Triangle", Journal Of Business Ethics, 2019.
6. Nadia Smaili; Paulina Arroyo; "Categorization Of Whistleblowers Using The Whistleblowi ng Triangle", Journal Of Business Ethics, 2019.
7. Noorul Azwin Binti Md Nasir; Muhammad Jahangir Ali; Kamran Ahmed; "Corporate Gov ernance, Board Ethnicity And Financial Statement Fraud: Evidence From Malaysia", Acco unting Research Journal, 2019.
8. Rr Maria Yulia Dwi Rengganis; Maria Mediatrix Ratna Sari; I. G. A. N. Budiasih; I Gde A ry Wirajaya; Herkulanus Bambang Suprasto; "The Fraud Diamond. Element In Detecting F inancial Statement Of Fraud", International Research Journal Of Management, It And Soci al Sciences, 2019.
9. Olubukola Ranti Uwuigbe; O Olorunshe; Uwalomwa Uwuigbe; E Ozordi; Osariemen Asiri uwa; T O Asaolu; Olayinka Erin; "Corporate Governance And Financial Statement Fraud Among Listed Firms In Nigeria", Iop Conference Series: Earth And Environmental Science, 2019.
10. Tomy Rizky Izzalqurny; Bambang Subroto; Abdul Ghofar; "Relationship Between Financi al Ratio And Financial Statement Fraud Risk Moderated By Auditor Quality", InternationalJournal Of Research In Business And Social Science (2147-4478), 2019.
11. Mousa Albashrawi; "Detecting Financial Fraud Using Data Mining Techniques: a Decade Review From 2004 To 2015", Journal Of Data Science, 2021. (If: 3)
12. Wei Dong; Stephen Shaoyi Liao; Zhongju Zhang; "Leveraging Financial Social Media Da ta For Corporate Fraud Detection", Journal Of Management Information Systems, 2018.
13. Dwi Ratmono; Yuvita Avrie Diany; Agus Purwanto; "Dapatkah Teori Fraud Triangle Me njelaskan Kecurangan Dalam Laporan Keuangan?", Jurnal Akuntansi Dan Auditing, 2018.
14. Yuehua Xu; Lin Zhang; Honghui Chen; "Board Age And Corporate Financial Fraud: an I nteractionist View", Long Range Planning, 2018.

15. Einde Evana; Mega Metalia; Edwin Mirfazli; Daniela Ventsislavova Georgieva; Istianing sih Sastrodiharjo; "Business Ethics In Providing Financial Statements: the Testing Of Frau d Pentagon Theory On The Manufacturing Sector In Indonesia", Business Ethics And Lead ership, 2019.
16. Vasant Raval; Vivek Raval; "Differentiating Risk Factors Of Ponzi From Non-Ponzi Frau ds", Journal Of Financial Crime, 2019.
17. Emily M. Homer; "Testing The Fraud Triangle: a Systematic Review", Journal Of Financi al Crime, 2019.
18. Charles Kusaya; John O'keefe; "Insider Abuse And Fraud Prediction For U.S. Banks: A C omparison Of Machine Learning Approaches", Accounting Technology & Information Sys tems Ejournal, 2020.
19. Razif Rosli; Intan Salwani Mohamed; Nafsiah Mohamed; Rohana Othman; Nabilah Rozz ani; "Development Of Fraud Prevention (Fp) Model Using The Theory Of Planned Behavi or", Business And Economic Research, 2020.
20. Lukas Brenner; Tobias Meyll; Oscar Stolper; Andreas Walter; "Consumer Fraud Victimiz ation And Financial Well-Being", Journal Of Economic Psychology , 2020.
21. Khaled Gubran Al-Hashedi; Pritheega Magalingam; "Financial Fraud Detection Applying Data Mining Techniques: a Comprehensive Review From 2009 To 2019 ", Comput. Sci. R ev., 2021.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of SOAP and/or the editor(s). SOAP and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.