M-score and F-score from the Financial Statement for Company Fraud Prediction

Hanyu Xu[1,a,*], Wenqi Shi[2,b], Jingyu Xia[3,c], Manchao Liu[4,d]

[1]*School of Arts and Science, Rutgers University, New Brunswick, 08901, US*
[2]*Lynbrook High School, San Jose, 95129, US*
[3]*Business School, Beijing Technology and Business University, Beijing, 100048, China*
[4]*Business School, McMaster University, Hamilton, L8s4l8, Canada*
a. hanyu.xu@rutgers.edu, b. wshi8785@gmail.com, c. 1465974395@qq.com, d.
lium106@mcmatser.ca
*corresponding author

Abstract: This paper investigates the circumstances where M-score and F-score are predicted to find out company fraud from financial statement. The first part goes over the background information for company fraud by answering three questions. The second part introduces the important parameters cited in the research and calculation. The third part shows the result by providing charts and analysis. The last part summarizes this paper. The importance of these two scores is hypothesized to increase the public ability to be conscious of company fraud from hidden data and false statements. However, under specific circumstances, M-score and F-score are not reliable reflections of other parameters like net income and accruals. The results of the data analysis are consistent with our prediction. This work shows the importance of individual thinking as a result does not agree with the accuracy of the model. One of the most significant creative thoughts in this work is about how to break through the traditional blind faith in established model when the result does not follow the standards.

Keywords: Fraud prediction, F-score, M-score

1. Introduction

1.1. What Is Financial Statement Fraud

Financial statement fraud means that deliberate change in the balances in a financial statement makes a company looks like have better operations and makes higher profits than it should be. Financial statement fraud often occurs accompanied by corporations' misrepresentation, falsifying balance sheets, and deceiving investors. Besides causing direct financial loss, fraud also damages a company's reputation. It may hamper the company's attraction to future business partners and investors. For listed companies in the US, fraud will make the company delist.

1.2. Why Managers Falsify Financial Statements to Mislead Investors

To answer the question, we have to figure out what the characters managers and investors play in the interaction of financial statements. Managers usually collect information about corporate accounting

and earnings reports to produce a financial statement and share it with directors, shareholders, auditors, and regulators. Investors are only a part of the readers of the financial statements who want to catch up on the newest information about the company they care about. Since they are only connected to each other by a document of numbers, it is not difficult to falsify the financial statement, and here are some basic reasons for that. Firstly, in order to meet future performance expectations and make the public trust in the ability of the company to make profits, managers have to report false revenue and fake income when a company cannot reach its goal in reality. Secondly, investors have nothing to do with detecting the manipulation because of the relationship between managers and auditors who are responsible for examining financial statements and making disclosures. Auditors would probably do something to make managers satisfied so they can benefit. Finally, due to the flexibility of the GAAP standards, it is feasible and easy to do the manipulation without exposure. After knowing about the reason that causes financial fraud, we will now move to the next part, which introduces mechanisms that prevent fraud.

1.3. What Are the Mechanisms That Deter Managers from Fraud?

There are four main mechanisms that works to deter managers from fraud.

Firstly, strengthen the internal moral education of enterprises. Managers are usually the initiator of accounting fraud. With high ethics, integrity, responsibility, and law-abiding of management and curb the greed of managers is conducive to eliminating the origin of fraud.

Secondly, improve internal control system and improve execution. Listed companies should strengthen the understanding of their own, timely find the loopholes, and combine with the actual situation to establish a sound internal control system. This can effectively monitor the performance of employees and strengthen the control of the accounting system. Management personnel should strictly divide the power and responsibility, and everyone should strictly follow the accounting rules and regulations to work to ensure true, reliable and comprehensive financial information. In addition, the internal audit department should also be strengthened, and fixed rules and regulations should be formed to match the size of the company.

Thirdly, improve the system of directors and supervisors. The board of directors of most listed companies in China now uses cumulative voting, which allows large shareholders to easily control the board. Changing the method of one person, one vote can safeguard the legitimate rights and interests of minority shareholders. Members of the Board of Supervisors should also have good professional knowledge, rich supervisory experience, lofty ethics, and professional ethics. Their institutional structure and scope of authority can learn from some advanced western experience. The proportion of independent supervisors in the board of supervisors should increase. Independent supervisors represent independence and fairness. The more independent supervisors are, the more difficult for the management to commit accounting fraud.

Finally, Strong external supervision puts pressure on managers. The media has a strong power to disseminate information and supervise by public opinion. As an important channel for the public to obtain information, the media can effectively improve the disadvantaged position of the public in information asymmetry. CSRC may consider issuing relevant policies and regulations to clarify the scope and intensity of media participation in supervision and clarify the responsibilities and powers of media supervision. As early as 1985, a number of associations in the United States jointly established the anti-fraud committee, while in China, it was established in 2015. Because it was established later and initiated by enterprises, there are problems such as fewer members, low influence and weak professionalism. In this regard, we can learn from the experience of the United States and get more professional associations and experts involved.

2. Model Description and Variable Measurement

2.1. Data

Our data come from the US listed Chinese company's annual financial report or statement. The first part has ten fraud companies that are (1) officially revealed by the regulatory authority, (2) charged by a short agency with financial fraud, and (3) delisted from the US list. The second part consists of 10 Chinese companies listed in US-listed without fraud accusations. Each part has 10 companies to make sure our calculations and results are reliable. For fraud companies, we use a four-year time period financial statement before the fraud appears, and for non-fraud companies, we obtain the recent four-year data. The following figure shows the twenty companies' names in each part.

Table 1: 10 Fraud and 10 Non-Fraud Chinese US listed companies used to calculate M-score.

Fraud (**M-score**)	not fraud
Luckin coffee	China National Petroleum Corporation
TAL Education Group	China Petroleum & Chemical Corporation
Ke Holdings Inc	China Life Insurance Company Limited
New Oriental Education & Technology Group Inc	Trip.com Group Ltd
Alibaba Group Holding Limited	bilibili
AutoChina International	NetEase
China energy corporation	Mining Limited
China Century Dragon Media	JD.com
IQIYI	Zhaopin Limited
Chemical	Thunder

Table 2: 10 Fraud and 10 Non-Fraud Chinese US listed companies used to calculate F-score.

Fraud (**F-score**)	not fraud
Luckin coffee	China National Petroleum Corporation
TAL Education Group	China Petroleum & Chemical Corporation
Ke Holdings Inc	China Life Insurance Company Limited
New Oriental Education & Technology Group Inc	Trip.com Group Ltd
Alibaba Group Holding Limited	Bilibili
China Intelligent Lighting and Electronics, Inc.	NetEase
China energy corporation	Mining Limited
China Century Dragon Media	JD.com
Physical Property Holdings Inc.	Zhaopin Limited
Keyuan Petrochemicals, Inc.	Thunder

After determining what companies are our research objectives, as shown in Table 1 and Table 2 we have to learn about what variables are necessary for model calculation. The next part will list our requirements of variables for two different models.

2.2. Variable

Some variables are representing a number of parameters in one year, which can be quoted directly, but other variables are meaning the change rate from the last year to this year, so we actually require five years data to calculate the change for specific variables.

2.2.1. M-score Variable

Receivables Index = DSRI
Gross Margin Index = GMI
Asset Quality Index = AQI
Sales Growth Index = SGI
Depreciation Index = DEPI
SG&A Expense Index = SGAI
Leverage Index = LVGI
Total Accruals to Total Assets = TATA

2.2.2. F-score Variable

Rsst_acc = ΔNon-cash net operating assets/Average total assets
Non-cash net operating assets=Stockholders' equity – Preferred stock – Cash & equivalents
Ch_rec = ΔReceivables/Average total assets
Ch_inv = ΔInventory/Average total assets
Soft_assets = (Total assets – PP&E, net – Cash & equivalents)/Total assets
ch_cs = % change in (Sales – ΔReceivables)
Ch_roa = Change in ratio of Net income/Average total assets
Issue = 1 if LTD debt or common and/or preferred equity issued
Net income is before extraordinary items (accidents or natural disasters) and discontinued operations [1].

2.3. Model

2.3.1. M-score

M-score is a mathematical model that M. Daniel Beneish created, and it is used to determine whether a company manipulates its earnings with financial ratios and eight variables. The eight variables are required for calculating M-score from the data of company's income statements, balance sheet and cash flow statements. They are Days Sales in Receivables Index (DSRI), Gross Margin Index (GMI), Asset Quality Index (AQI), Sales Growth Index (SGI), Depreciation Index (DEPI), SG&A Expense Index (SGAI), Leverage Index (LVGI), Total Accruals to Total Assets (TATA).

The basic theory that Daniel Beneish based on is(1) year over year sales is growing quickly and extremely high.(2)a decline in asset quality, profit margins and increase in leverage.(3) using aggressive accounting practices(receivables growing much faster than sales; decreasing depreciation expense; a large amount in increasing accruals.) The calculated M-score can show the degree of manipulations in earnings and higher Beneish M-score means higher probabilities of manipulating earnings. If Beneish M-score is less than -2.22, the company is under consideration of being a manipulator. If the Beneish M-score is greater than -2.22, it shows that the company can be a manipulator. [2]

There was an example which in 1998, a group of Cornell University students used the Beneish model to predict that Enroll Corporation was manipulating their earnings. At that time, the stock price

of Enroll was 48 dollars per share and it was eventually climbed to 90 dollars per share before it bankrupted in 2001. The Cornell students did sound the alarm, but nobody heard their advice. For the M-score model, we are going to calculate the M-score with the data of financial statements (income statements, balance sheet, cash flow statements) in the sample of 20 Chinese companies that listed in the USA by M-score formula and analyze the results of the M-score then see whether some Chinese companies are able to manipulate in their earnings.

Here is the formula of M-score:

$$M - score = -4.840 + 0.920 * DSRI + 0.528 * GMI + 0.404 * AQI + 0.892 * SGI + 0.115 * DEPI - 0.172 * SGAI - 0.327 * LVGI\ 4.679 * TATA$$

2.3.2. F-score

Dechow F-score is a seven-variable model designed to identify companies' probability of financial misstatement. The seven variables refer respectively to Change in non-cash net operating assets (rsst_acc), Change in receivables (ch_rec), Change in inventory (ch_inv), Percentage soft assets (soft_assets), Change in cash sales (ch_cs), Change in return on assets (ch_roa), Debt or equity issuance (issue) [1].

Dechow et al. analyzed 2,190 AAERs (Accounting and Auditing Enforcement Releases companies) released (by the U.S. Securities and Exchange Commission) between 1982 and 2005, in which they identified 676 firms that had fraud on their quarterly or annual financial statements. Dechow compared a variety of features of those fraud companies and non-fraud companies which included accrual quality, financial performance, nonfinancial performance, off-balance-sheet activities, and market-related variables (such as inventory and the rate to employ or fire employees) [1].

For the computation of the Dechow F-score, the first step is to calculate the predicted value with the formula Predicted value = -7.893 + 0.790*rsst_acc + 2.518*ch_rec+ 1.191*ch_inv + 1.979*soft_assets + 0.171*ch_cs − 0.932*ch_roa + 1.029*issue. The higher the predicted value calculated above, the higher probability of manipulation. For the next step, we use logistic function •=e^Predicted value/(1 + e^Predicted value), where e = 2.71828183, to calculate probability of manipulation; the last step is calculating the F-score by dividing the probability of manipulation by unconditional probability of manipulation (0.0037) [1]. The smaller the F-score is, the lower the probability of a company misstating its financial statement; in contrast, the higher it is the higher the probability. The Watersheds of interpretation the F-score are commonly 1 and 2.45.

A F-score smaller or equal to 1 indicates the risk of a company misstating its financial statements is very low, whereas F-score greater 1 indicates "above normal risk" and F-score greater than 2.45 indicates "high risk". According to Dechow et al. (2011)'s data, 63.7% of non-misstated companies in his research demonstrate a F-score smaller than 1 and about only 5-6% of them demonstrate a F score greater than 2.45. On the other hand, only 31.4% of the misstated companies have a F-score smaller than 1 while about 18% percent of them have F-score greater than 2.45. Dechow F-score's mechanism mainly depends on different variables' numerical value, fluctuation and ratio [1]. For instance, Hung et al. examined the relationship between all the F-score variables and their findings demonstrated that rsst_acc, ch-rec, and soft assets all exhibit a strongly Positive Proportion with the probability of fraud. Each of the 7 variables can independently serve as an indicator to identify the feature of fraud, and then the mathematical metric balances each indicator's interpretation. The output of certain mechanism is F-score, which maximizes the accuracy of the model [2-3]; Aghghaleh reveals that Dechow F-score has an average accuracy of 76.22% in its ability to detect the financial fraud of companies, which is even higher than its accuracy of classical model Benish M-score which has an average accuracy of 73.17% [3-4]. Today our research group will use the F-score model to measure the financial performance of a sample of the U.S.-listed Chinese companies that committed

fraud and a sample of certain companies not fraud in the previous 10 years to find out the potential hints that would help investors and auditors to discover the risk of financial fraud sooner.

2.3.3. Accruals

As a first step, we know that a positive accrual ratio indicates that earnings are less persistent and depend more on accruals, while a negative accrual ratio indicates that earnings are more persistent and depend more on cash. According to the formula (net income $12 million - cash from operations $12 million - cash from investments $12 million) / (total assets in the first quarter of this year) + (total assets in the first quarter of last year), the accrual ratio was calculated by comparing 10 fraud companies to 10 non-fraud companies. The results of the data can be analyzed in several ways. An example is inventory management. Cash spent on inventory purchases will decrease a company's liquidity if it stocks up on inventory. Assuming the purchase is used to fulfill current sales based on existing inventory, current earnings will not be affected. In order to benefit from future economic growth, the company increases its inventory. Depending on whether the inventory is sold above or below its cost, actual future earnings will vary. For the purpose of this discussion, we will use China Intelligent Lighting and Electronics as an example. Consider the case where the company buys $20,000 of lights in the first quarter because it believes the product will sell well in the second quarter, which leads to an increase in inventory, but profits remain the same. During the first quarter, the company spent $20,000 on the purchase. In the second quarter, inventory forecasted to be sold was not as good as expected for some reason. Despite remaining unsold inventory, the company was determined to mark down the remaining inventory to half of its cost. Even though all the lights were sold in the end, the $10,000 charge was included in the profit during the second quarter. Cash flow would differ from earnings in the first quarter. Based on this example, the company spent $20,000 on lights, reducing cash flow, but not first quarter cash flow. The transaction results in a cash outlay, but no profits or sales. Therefore, earnings would exceed operating cash flow. Because of write-downs, earnings are lower in the second quarter of the second year, but cash flow is unchanged. Compared with accrual-based earnings, cash earnings persist. The company's optimism will result in write-downs resulting in less persistent earnings attributable to accruals. Future profitability will be reduced by these write-downs. Accounting motivation can be measured by accruals. As a result of a high level of accruals, the company will be notified of any upcoming write-downs in advance.

2.4. Calculation

To get each variable for model calculation, we first put collected data into Excel for the pretreatment. For the F-scores, we follow the instruction from above to get the predicted value. Basically, we calculate the first company data and use the autofill to get others. The next step is to calculate the probability of manipulation, which is the same as the previous step with autofill. Finally, to get the F-score, we follow the formula as well. To be honest, the calculation part for F-score is not hard work for us compared to the data collection and input, which are extremely repetitive and laborious. For the M-scores, we decide to use an online calculator since there are too many variables.

2.5. Accruals

Accrual accounting, also known as the "receivable and payable system", determines whether expenses and revenues incurred in the current accounting period should be recognized in the current period's profit or loss. If monetary funds were received or not, all expenses incurred in the current period should be reimbursable from the income of that period; whether monetary funds were actually received or not, any income received in the current period should be treated as income for that period. Through the implementation of this system, the level of expenses and profit and loss of each period

are accurately reflected in Net income 12m – cash from operations 12m – cash from investing 12m/(total assets Q1 × 0.5) + (total assets Q5 × 0.5)×100.

How accrual works: According to the American Association of Individual Investors, the authors describe fictional food truck businesses in two examples. On the first day, Bob rents a food truck for $1,000 and purchases $1,000 worth of items. It realized $2,500 in revenue when everything was sold out on the first day. Due to his $2000 investment in the food truck and other items, he will make $500 profit on that day. The $500 is then the profit in cash. (2500 - 2000 = 500) Sue's operation, however, uses the accrual system. Having high hopes of success, she invests $30,000 on a 30-day food truck. After 30 days, she planned to buy a better food truck with her profits. Despite earning $2,500 on the first day, she wasn't set up to accept credit cards. Because of this, Sue told half of the customers without cash that they could pay with a credit card or in cash the next day. As a result of using the accrual system, Sue earned the same profit on the first day as Bob. In this example, Sue's food truck is counted to depreciate at $1,000 per day because it can only be used for 30 days before it needs to be replaced. Both Sue and Bob's expenses would total $2,000 due to Sue's $1,000 food expense. The two problems in Sue's operation are related to the fact that half of her customers do not pay her cash on the day, which means it is possible that she will not be paid. Additionally, she needs to ensure that she makes $500 a day in profit. Her assumption was just optimistic. According to Sue's balance sheet, her accounts receivable total $1,250. The money she owes her is from food purchases made by customers. In accrual accounting, Sue can record the full value of her first-day sales on her income statement as long as she lists the money owed to her as receivable in her current assets section. Due to the fact that Sue has received payment from her customers, Sue's cash flow statement reduces cash generated from operating activities by $1,250. According to Sue's balance sheet, the truck is worth $30,000. In the statement of cash flows, she recorded an outflow of $30,000 as a result of the cash outlay for fixed assets. The line on the cash flow statement represents cash from investing activities. On day one, $1,000 is recorded as a depreciation expense to account for the loss of value of the truck due to depreciation. On the balance sheet, the truck's value is adjusted by subtracting the depreciation expense. Consequently, the profit is reduced by $1,000 due to the decrease in the value of the asset, which is a cost. This expense, however, was not paid in cash by Sue. The statement of cash flows adds depreciation expenditures to cash from operating activities. As a result of all the accounting, Sue's free cash flow is -29,750 ($31,000 in cash expenses minus $1,250 in reimbursable transactions)

3. Results

3.1. Charts

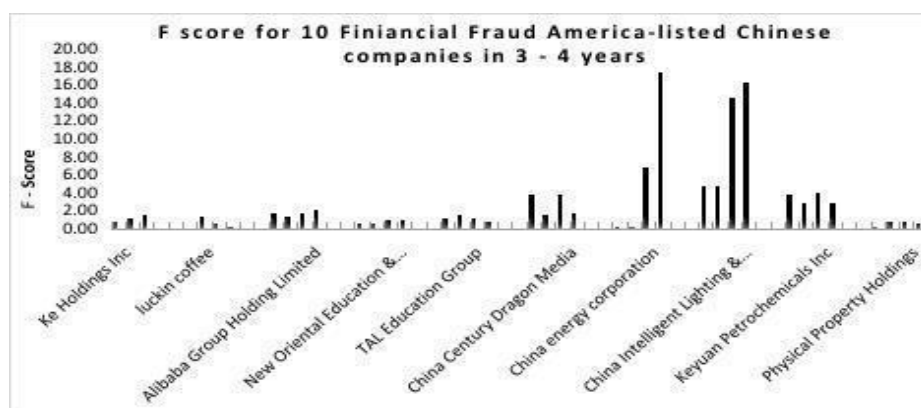#### 3.1.1. F-Score Charts (Excluded 2 Outliers Which Are > Than 100 from Graph)

Figure 1: F score for 10 financial fraud America-listed Chinese companies in 3-4 years.

Figure 2: F score for 10 financial non-fraud America-listed Chinese companies in 3-4 years.

3.1.2. M-score Charts

Figure 3: M score for 10 financial non-fraud America-listed Chinese companies in 3 years.

Figure 4: M score for 10 financial fraud America-listed Chinese companies in 3 years.

3.2. Analyze

3.2.1. F-score Analysis

Overview of result: The Figure 1 and 2 above exhibit the result of F-score of 10 Fraud and 10 Non-Fraud US. Listed Chinese companies respectively. The results do not fit our expectations before the

calculation. Based on the interpretation of Dechow F-score, we conjectured that the F-score of fraud companies would be significantly higher than those of non-fraud companies, and the closest to the Fraud year, the obvious this trend would be. However, neither hypothesis 1 nor 2 cannot be verified with the results of calculation. The only signal we can figure out from our data is that the fraud set companies are more typically to have a much greater fluctuation in F-score before 2 - 3 years of financial misstatements; on the other hand, non-fraud companies typically exhibit more stability, no matter the value of their F-score are big or small. (For an individual company)

Reason for divergence of our expectations and the results: Dechow F-score has two innate weaknesses: One is from the mechanism of the calculation, while the other is from the limits of Dechow's research scope. The F-score reflects the potent strength of each variable in analyzing a company's financial performance. However, one number is assigned to the first feature, and another to the second: it makes no mention of the combination of these variables (mutual function), which is F-primary score's flaw [5].

The second point is that Dechow undertook a thorough examination of 2,190 AAERs accessible between 1982 and 2005 in order to identify enterprises with incorrect quarterly or yearly earnings; the limit here is that all of those companies are AAERs accessible, and the time was between 1982 and 2005 [1]. A few decades have passed, and the same model may not be that fit to modern companies because how companies operate their business today might be quite different from those of last century. The third thing is that the sample size of our data is really small and variable: we have only 10 companies for each group and almost all of them come from different industries; this makes it unlikely to investigate the relationship between numerical value of F-score and Fraud/Non-Fraud.

The calculation of F-score gives us a potential signal that can be further investigated and might help investors and auditors better to identify if a company misstates its financial performance: Fluctuation of F-score in a few years for the same company. The greater fluctuation might exhibit a greater risk of financial misstatement and there would be some key values that can help us to identify what means high risk, normal risk, and low risk. This is a potential direction for further research, and it requires a much greater sample size.

3.2.2. M-score Analysis

The calculated results of M-score for two groups of companies (fraud and not fraud), as shown in Figure 3 and 4 above respectively, are relatively in line with our expectations. Through our understanding and interpretation of m score, Figure 3 shows that m-score is common for not fraud companies and that the result will be less than -2.2. Figure 4 shows that for fraud companies, the results will generally be greater than -2.2. And the larger the m-score result, the more likely the company is to manipulate earnings. Fraud's m-score value will be higher than that of not fraud company and the value is larger. Also, the closer to the fraud year, the more obvious the increase in M-score for fraud. Based on the calculated data, it can be found that the results are more in line with the above expectations.

The main function of M-score is to predict the possibility of a company manipulating earnings in the future, and it can also release risk signals to investors through the data of financial statements, helping them make better investment decisions. Secondly, it can also help investors, auditors, and regulators to better avoid risks and be used as evidence of financial fraud.

4. Conclusion

This paper hypothesizes that there exist parameters from financial statements which can be used to provide a measure of company fraud than obvious data like net income and accruals. Our result is consistent with our prediction. First, many reasons behind financial statements make it hard for

investors to distinguish whether a company is a fraud or not even though there are some deter mechanisms to provide protection. Second, simple data which can be directly obtained from the financial statements is cheated, so it is still difficult to figure out the fact based on net income and others. Finally, we decided to find out other measurements which are able to reflect the real situation of a company falsification——F-score and M-score. F-score and M-score are closely related to fraud, and they are reliable measurements and provide enough information to investors. In terms of these conditions, we hypothesize that these two scores are more relevant to determine whether a company is fraudulent or not. This paper also provides insights into accruals and their importance to identify fraud. Through the implementation of the Accruals system, the level of expenses and profit and loss of each period accurately reflect Net income, so that the calculation would be clearer and more convenient.

The contribution comes out of this paper is to analyze the feasibility of applying F-scores and M-scores to financial fraud. Specifically, this paper explains why we choose F-score and M-score as more reliable resources compared to net income and accruals. We assume that these two scores are related to more information like Receivables Index, Gross Margin Index, Average total assets, and Change in ratio of Net income, elements that can fully reflect the real situation of a company operating in a long time period, where managers are hard to do data fraud. Even though purely net income or assets are meaningless in the evaluation of fraud, it is not saying that we should exclude these basic elements from our calculations, and they should be considered together with others and make a conclusion. Another instructive advice comes from the calculation part; we have to confirm that there is heavy work for data collection and input by human effort, it would be better to be programmed in order to perform huge and repetitive tasks.

The passage shows that there exist some methods that work for the public to detect fraud in a company from accessible resources and provide the probability that by analyzing a long time period of data, we can find what managers want to hide in financial statements since they can only change one-year statements rather than all of the financial statements. And we also acknowledge that in specific situations, M-score and F-score cannot fully reflect the situation. Finally, future research on the same topic can investigate additional equations and calculations to get more exact results since there is still an outlier in this paper.

References

[1] DECHOW, P. M., GE, W., LARSON, C. R. and SLOAN, R. G. (2011), Predicting Material Accounting Misstatements. Contemporary Accounting Research, 28:34-38

[2] Hung, D. N., Ha, H. T. V., & Binh, D. T. 2017. Application of F-Score in Predicting Fraud, Errors: Experimental Research in Vietnam. International Journal of Accounting and Financial Reporting 7(2): 303-322

[3] Tahani Ali Hakami, Mohd Mohid Rahmat, Mohd Hasimi Yaacob & Norman Mohd Saleh(2020). Fraud Detection Gap between Auditor and Fraud Detection Models: Evidence from Gulf Cooperation Council.Asian Journal of Accounting and Governance 13: 1–13

[4] Aghghaleh, S. F., Mohamed, Z. M., & Rahmat, M. M. 2016. Detecting Financial Statement Frauds in Malaysia: Comparing the Abilities of Beneish and Dechow Models. Asian Journal of Accounting and Governance 7: 57-65.

[5] StackExchange,2016.The disadvantage of using F-score in feature selection. https://stats.stackexchange.com/questions/20341/the-disadvantage-of-using-f-score-in-feature-selection.