Available online at www.ijournalse.org

**Emerging Science Journal**

(ISSN: 2610-9182)

Vol. 6, No. 6, December, 2022

# Discovering Future Earnings Patterns through FP-Growth and ECLAT Algorithms with Optimized Discretization

Putthiporn Thanathamathee [1], Siriporn Sawangarreerak [2*]

[1] *School of Engineering and Technology, Walailak University, Nakhon Si Thammarat 80160, Thailand.*

[2] *School of Accountancy and Finance, Walailak University, Nakhon Si Thammarat 80160, Thailand.*

**Abstract**

Future earnings indicate whether the trend of earnings is increasing or decreasing in the future of a business. It is beneficial to investors and users in the analysis and planning of investments. Consequently, this study aimed to identify future earnings patterns from financial statements on the Stock Exchange of Thailand. We proposed a novel approach based on FP-Growth and ECLAT algorithms with optimized discretization to identify associated future earnings patterns. The patterns are easy to use and interpret for the co-occurrence of associated future earnings patterns that differ from other studies that have only predicted earnings or analyzed the earnings factor from accounting descriptors. We found four strongly associated increases in earnings patterns and nine strongly associated decreases. Moreover, we also established ten accounting descriptors related to earnings: 1) %Δ in long-term debt, 2) %Δ in debt-to-equity ratio, 3) %Δ in depreciation/plant assets, 4) %Δ in operating income/total assets, 5) %Δ in working capital/total assets, 6) debt-to-equity ratio, 7) issuance of long-term debt as a percentage of total long-term debt, 8) long-term debt to equity, 9) repayment of long-term debt as a percentage of total long-term debt, and 10) return on closing equity.

**Keywords:**

Future Earnings Patterns;
Financial Statement;
Association Rule Mining;
Optimized Discretization.

**Article History:**

| | | | |
|---|---|---|---|
| **Received:** | 27 | May | 2022 |
| **Revised:** | 29 | July | 2022 |
| **Accepted:** | 18 | August | 2022 |
| **Available online:** | 13 | September | 2022 |

## 1- Introduction

At present, the decisions of investors or users of financial statements are based on information published in a business entity's financial statements for analysis to assess the company's valuation. The present analysis results are frequently used to make decisions for investors. For creditors, the analysis results are employed to consider the sustainability and continuation of business operations, including credit granting or credit in joint trade. Forecasting future earnings is helpful for financial analysts in making investment recommendations. Investors use this to make investment decisions in businesses with consistently good performance, while creditors want to know the possibility of obtaining repayment from a business entity to honor its contractual obligation. In addition, the organizational executives need to plan and adjust the future organizational strategy. Consequently, a model for forecasting future earnings would significantly help determine how to formulate organizational development strategies and plan toward achieving its targets. As a result, financial analysts, creditors, investors, and corporate executives all wish to predict the future earnings of an organization. To this end, the underlying data used in the forecast can be cash flow, profit, or other accounting components [1].

For most decision-making, the financial statement's users apply earnings analysis to look for trends in an entity's future performance. The profit and loss are determined from the revenues and expenses incurred in the accounting period [2, 3]. If a business has more revenue than expenses, it will affect its profit. However, if business revenues are lower than expenses, this will affect its loss. Previous studies have shown that forecasting future profits using past profits as a

* **CONTACT**: siriporn.sa@wu.ac.th

**DOI:** http://dx.doi.org/10.28991/ESJ-2022-06-06-07

predictor is more effective than considering cash flow [4, 5]. Daraghma (2013) [5] has found that the previous year's earnings have an essential role in predicting future earnings, whereas the previous year's operating cash flows are irrelevant. However, some previous works have also found that forecasting future profits using cash flow is a better predictor than past earnings [6, 7]. Arther et al. (2010) [6] have proposed that the cash flow components model is superior to an aggregate cash flow model in terms of explanatory power and predictive ability for future earnings. Shubita (2021) [7] has shown that the cash flows from operations, investing, and financial activities statistically impact predicting future earnings.

Most studies on earnings prediction employ statistical techniques. For example, Ball and Watts (1972) applied random walk and time-series models to predict future earnings [8]. Ou & Pennman (1989) have proposed an earnings prediction model with financial statement data [9]. They developed their prediction model according to the direction of future change (positive change or negative change) in earnings using stepwise logistic regression. Ball & Ghysels (2017) have developed a mixed data sampling (MIDAS) regression method to forecast firm-level earnings [10]. Hou et al. (2012) have proposed a new approach to estimating the implied cost of capital (ICC) using the earnings forecasts from a cross-sectional model (HVZ model) [11]. Gerakos & Gramacy (2013) have developed a simple Random Walk (RW) model as an earnings forecasting model [12]. Li & Mohanram (2014) have developed the Residual Income (RI) model or Earnings Persistence (EP) model to predict earnings [13]. Both models outperform the previous HVZ model. Azevedo et al. (2021) have proposed a method to forecast corporate earnings [14] by employing a parsimonious cross-sectional regression model consisting of analyst forecasts of earnings, gross profits, and past stock performance, leading to a more accurate and less biased result.

Nowadays, the number of financial statements is high, making statistical techniques inadequate to classify data correctly. Therefore, machine learning techniques must be applied to analyze big data to predict future earnings. Anand et al. (2019) [15] applied the classification random forest algorithm to predict directional changes (increase or decrease) in return on equity (ROE), return on assets (ROA), return on net operating assets (RNOA), cash flow from operations (CFO), and free cash flow (FCF). The classification accuracy of this method is better than the Random Walk (RW) model. Hunt et al. (2019) have modified the stepwise logistic regression with an elastic net to improve the prediction accuracy of future earnings [16]. Nguyen (2020) used gradient-boosted regression trees to forecast earnings. The latter study can outperform analysts in their earnings prediction [17]. Chen (2022) used decision trees, random forests, and stochastic gradient boosting methods with high-dimensional detailed financial data to predict the direction of one-year-ahead earnings changes [18]. These models significantly outperform the stepwise logistic regression proposed by Ou and Penman's work [9]. Ishibashi et al. (2016) focused on variable selection for establishing an earnings prediction model [19] based on the same 68 variables in [9]. This work applied relief, correlation-based feature selection, and consistency-based subset evaluation as feature selection techniques and used C4.5 decision trees as a learner. Then, the prediction model was constructed by logistic regression. The results revealed that about 10-23 variables influenced the future earnings prediction.

From the previous studies, regression and other classification modeling have gained popularity in predicting future earnings to assess the statistical hypothesis interactions among independent variables or find important accounting descriptors that affect earnings. Nevertheless, it is difficult to interpret and analyze the co-occurrence of accounting descriptors. As a financial analysis system generates large data containing hidden knowledge, identifying associated future earnings patterns is deemed impossible from previous works. According to the literature review, the following research questions were proposed in the present study:

***RQ1:*** *What machine learning techniques can be used to identify future earnings patterns in financial statements?*

***RQ2:*** *What associated patterns and accounting descriptors are related to future increased and decreased earnings after finding future earnings patterns?*

This study aims to make the following contributions based on previous research and the above research questions. First, FP-Growth and Eclat association rule mining is applied to identify the strongly associated future earnings patterns. This is a new way of considering future increased and decreased earnings, which differs from research that looks at individual factors. The results of the associated patterns give users of this approach a clear and easy-to-understand relationship with the future earnings increase or decrease because associated future earnings patterns identify correlated factors and specify the range of co-occurrences of the factors. This is different from other research that only finds factors that affect future earnings or only forecast future earnings. Second, to identify data ranges before finding associated future earnings patterns, we apply the optimization discretization technique to find the optimal range for the dataset. This is unlike previous research [20], in which users must specify the number of bins before converting continuous data attributes into discrete data attributes.

The remainder of this paper is organized as follows: Section 2 describes the theoretical background. The proposed method is discussed in Section 3. The experimental setup and results are presented in Section 4. Next, a discussion is presented in Section 5. Finally, the conclusions are presented in Section 6.

## 2- The Theoretical Background

### *2-1- Future Earnings*

Future earnings predictions using financial statement information is a profit prediction that will occur in the future of a business. It uses data in the financial statements to predict whether a business will have profits or losses in the future. Ou and Penman (1989) [9] have estimated the probability of earnings increase or decrease in the subsequent fiscal year by financial descriptors in the statements. Ishibashi et al. (2016) [19] have shown the current profit-earning per share for a given firm $i$ in the fiscal year $t$ as $e.p.s$ and identify $Pr_i(t+1)$ as the relative ability of firm $i$ to generate earnings in the subsequent fiscal year as in the equation:

$$Pr_i(t+1) = \begin{cases} 0 & (e.p.s_i(t+1) - e.p.s_i(t) - drift_i(t+1) < 0) \\ 1 & (e.p.s_i(t+1) - e.p.s_i(t) - drift_i(t+1) > 0) \end{cases}, \quad (1)$$

where $drift_i(t+1)$ is the average change of $e.p.s$ in the last four years from fiscal year $t+1$. If $Pr_i(t+1)$ is defined as an earnings increase of a firm, then $Pr_i(t+1) = 0$ is defined as an earnings decrease. Sixty-eight descriptors in the financial statements were derived from [9], as given in Table 1.

**Table 1. The 68 descriptors in the financial statements**

| No. | Accounting descriptor [a] |
|---|---|
| 1 | Current ratio |
| 2 | %Δ in current ratio |
| 3 | Quick ratio |
| 4 | %Δ in quick ratio |
| 5 | Days sales in accounts receivable |
| 6 | %Δ in days sales in accounts receivable |
| 7 | Inventory turnover |
| 8 | %Δ in inventory turnover |
| 9 | Inventory/total assets |
| 10 | %Δ in inventory/total assets |
| 11 | %Δ in inventory |
| 12 | %Δ in sales |
| 13 | %Δ in depreciation |
| 14 | Δ in dividend per share |
| 15 | Depreciation/plant assets |
| 16 | %Δ in depreciation/plant assets |
| 17 | Return on opening equity |
| 18 | Δ in return on opening equity |
| 19 | %Δ in (capital expenditure/total assets) |
| 20 | %Δ in (capital expenditure/total assets), one-year lag |
| 21 | Debt-equity ratio |
| 22 | %Δ in debt-equity ratio |
| 23 | Long term debt to equity |
| 24 | %Δ in long term debt to equity |
| 25 | Equity to fixed assets |
| 26 | %Δ in equity to fixed assets |
| 27 | Times interest earned |
| 28 | %Δ in times interest earned |
| 29 | Sales/total assets |
| 30 | %Δ in sales/total assets |
| 31 | Return on total assets |
| 32 | Return on closing equity |
| 33 | Gross margin ratio |
| 34 | %Δ in gross margin ratio |
| 35 | Operating profit (before depreciation) to sales |
| 36 | %Δ in operating profit (before depreciation) to sales |

| No. | Accounting descriptor [a] |
|---|---|
| 37 | Pretax income to sales |
| 38 | %Δ in pretax income to sales |
| 39 | Net profit margin |
| 40 | %Δ in net profit margin |
| 41 | Sales to total cash |
| 42 | Sales to accounts receivable |
| 43 | Sales to inventory |
| 44 | %Δ in sales to inventory |
| 45 | Sales to working capital |
| 46 | %Δ in sales to working capital |
| 47 | Sales to fixed assets |
| 48 | %Δ in production |
| 49 | %Δ in R&D |
| 50 | %Δ in (R&D/sales) |
| 51 | %Δ in advertising expense |
| 52 | %Δ in (advertising/sales) |
| 53 | %Δ in total assets |
| 54 | Cash flow to total debt |
| 55 | Working capital/total assets |
| 56 | %Δ in working capital/total assets |
| 57 | Operating income/total assets |
| 58 | %Δ in operating income/total assets |
| 59 | %Δ in total uses of funds |
| 60 | %Δ in total sources of funds |
| 61 | Repayment of long term debt as % of total long term debt |
| 62 | Issuance of long term debt as % of total long term debt |
| 63 | Purchase of treasury stock as % of stock |
| 64 | %Δ in funds |
| 65 | %Δ in long term debt |
| 66 | Cash dividend as % of cash flows |
| 67 | %Δ in working capital |
| 68 | Net income over cash flows |

[a] Δ indicates change. In calculating %Δ, observations with zero denominators are excluded and absolute values are used in all denominators [9].

### *2-2- Association Rule Mining*

Association rule mining is an unsupervised learning technique, an associated dataset pattern that appears in a database [21]. In the association rule mining, "if" represents a condition, and "then" represents a conclusion. There are three steps in finding a relationship rule [22] as follows:

- Support value is the percentage of datasets that appear in a database, which can be calculated as follows:

$$Support(A \rightarrow B) = P(A \cap B), \quad (2)$$

where, A and B is the dataset, $Support(A \rightarrow B)$ is the support from A to B, and $P(A \cap B)$ is the probability that B occurs when A occurs within the datasets.

- Confidence value is the percentage of the rule found. When A appears, then B appears at the same time. It can be calculated as follows:

$$Confidence(A \rightarrow B) = \frac{P(A \cap B)}{P(A)} \quad (3)$$

- Lift value is the percentage of the number of rule datasets present in a database to the number of datasets A and B. It can be as follows:

$$Life(A \rightarrow B) = \frac{P(A \cap B)}{P(A) \times P(B)} \quad (4)$$

If the $lift < 1$, $A$ and $B$ have a negative relationship.

If the $lift < 1$, $A$ and $B$ have a positive relationship.

If the $lift < 1$, $A$ and $B$ are independent.

### *2-3- Frequent Pattern Growth (FP-Growth) Algorithm*

Han et al. (2004) [23] have developed an algorithm to reduce the number of readings from a database. The FP-Growth algorithm reads the data from a database only two times and does not create a candidate dataset. It takes less time and memory to process than the Apriori algorithm [24]. The principle of the FP-Growth algorithm has two steps [20]:

- Creating a tree of frequent patterns by reading the data from a database twice. The first reading is to count the list of frequencies of a dataset. Then, take the dataset with a frequency greater than the minimum support value. Next, sort in descending order to create a header table. The next step is to read the data for the second time from the database to create an FP-Tree by reading each transaction data. Then, only the datasets that appear in the header table are selected to create a node in the FP-Tree, and each node is connected with the same dataset in the header table.
- Find a set of datasets that frequently occur together. Create a dataset that appears with the dataset being considered in each conditional pattern base and an FT-Tree on a conditional pattern base (Conditional FP-Tree). Consider the dataset in a bottom-up approach of the header table and select a Conditional FP-Tree with a frequency above the specified minimum support value. Then, find the common datasets between the Conditional Pattern Base and the Conditional FP-Tree of each dataset.

### *2-4- Equivalence Class Clustering and bottom-up Lattice Traversal (Eclat) Algorithm*

Zaki (2000) [25] has developed the Eclat algorithm that only scans the dataset once. Unlike Apriori and FP-Growth algorithms, the Eclat Algorithm uses a vertical dataset and depth-first search approach. The only support value is calculated in the Eclat algorithm. After calculating the support value of all items, they are compared with the minimum support value. The items with a support value above or equal to the minimum support value are generated in frequent itemsets. Zhang et al. (2021) [24] have suggested that the Eclat algorithm requires less time and is thus more efficient. However, as this algorithm needs to save itemsets repeatedly, it needs more memory space. The principle of the Eclat algorithm is shown in Algorithm 1.

**Algorithm 1.** Eclat Algorithm

**Input:** $F_k = \{I_1, I_2, \ldots, I_n\}$ //cluster of frequent $k$-itemsets.
**Output:** The complete set of frequent $l$-itemsets.
**Method:** Call Bottom-Up ($F_k$)
Procedure Bottom-Up ($F_k$)
{
(1) **for all** $I_i \in F_k$
(2) $F_k + 1 = \emptyset$
(3) **for all** $I_j \in F_k, i < j$
(4) $N = I_i \cap I_j$
(5) **if** $N.sup_1 \geq min\_sup$ **then**
(6) $F_k + 1 = F_k + 1 \cup N$
(7) **end**
(8) **end**
(9) **end**
(10) **if** $F_k + 1 \neq \emptyset$ **then**
(11) Call Bottom-Up ($F_k + 1$)
(12) **end**
}

### *2-5- Optimized Discretization*

The optimal binning is the optimal discretization of a variable into bins given a discrete or continuous numeric target. This work uses the OptBinning library developed by Navas-Palencia (2020) [26], a library written in python programming formulation to handle the optimal binning problem for a binary target. The optimal binning process has two steps: 1) a pre-binning process that generates an initial granular discretization, and 2) a subsequent optimization to satisfy the constraints. Navas-Palencia (2020) [26] has shown that an excellent binning algorithm for the binary target should be specialized by the following properties: 1) the missing values are binned separately, 2) each bin should contain at least 5% observations, and 3) no bins should have 0 accounts for non-events or events.

## 3- Proposed Methodology

The proposed method for identifying associated future earnings patterns is shown in Figure 1. It consists of three main steps: 1) data pre-processing and optimized discretization, 2) generating associated patterns using FP-Growth and Eclat algorithms, and 3) identifying the associated future earnings patterns.

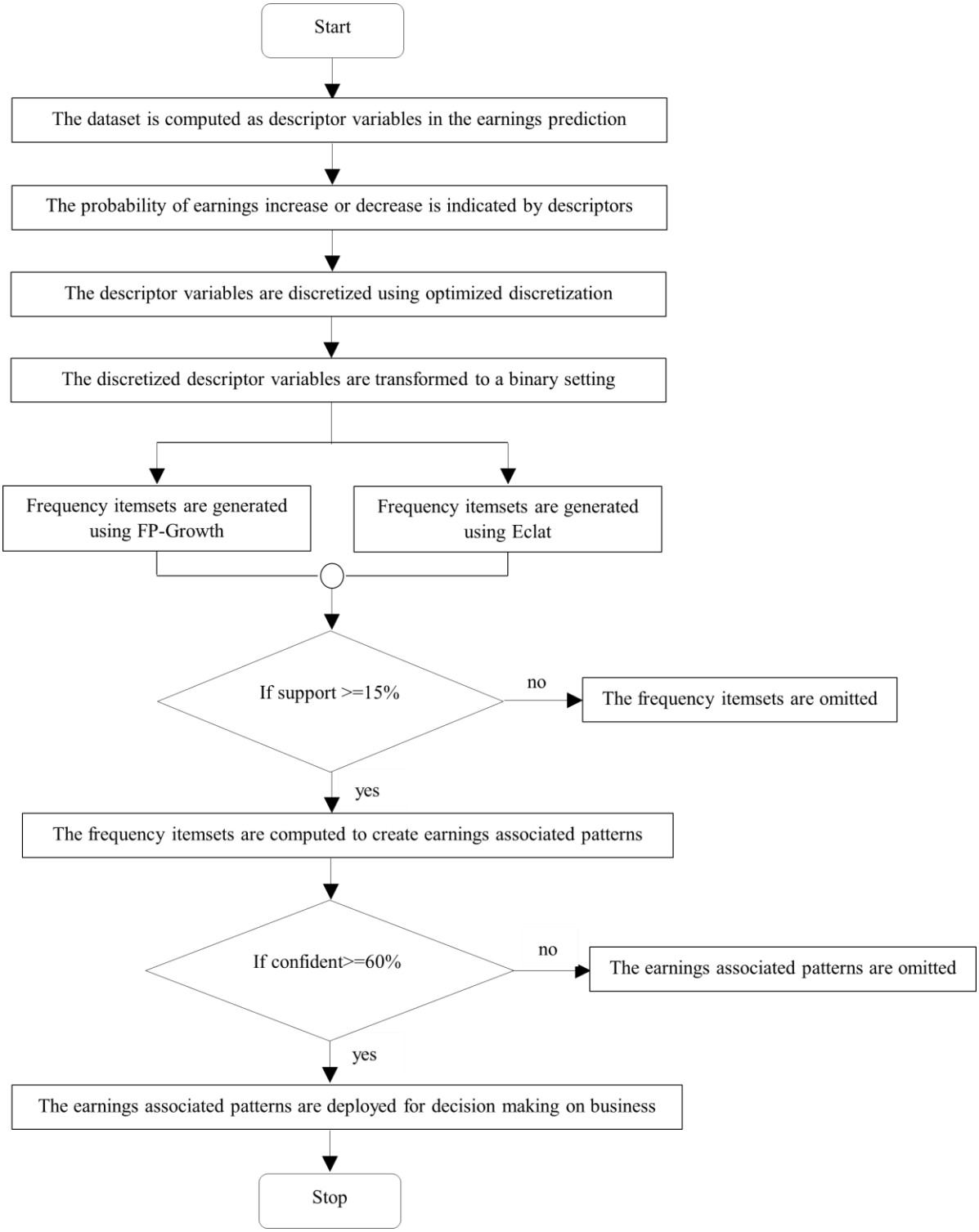


**Figure 1. Proposed method for identifying future earnings patterns**

### *3-1- Data Pre-Processing and Optimized Discretization*

A total of 2850 financial statements of 570 companies listed on the Stock Exchange of Thailand from 2016 to 2020 were computed using 68 accounting descriptor variables given in Table 1. Then, the probability of an increase or decrease in earnings in the subsequent fiscal year $Pr_i(t+1)$ was calculated as described in Section 2.1. Before generating association rules from the numerical data, 68 accounting descriptor variables were converted into discrete data attributes

using the optimized discretization technique. In this study, the technique finds the optimal bins without determining by the user [26]. The technique assigns an optimal bin for each attribute, which is better than binning with equal width or binning with equal frequency, whereby the user determines how many layers to divide the data into until the appropriate bin is found.

### *3-2- Generating Associated Patterns using FP-Growth and Eclat Algorithms*

Association rule mining was used to find associated patterns. This study used two association rule mining algorithms: FP-Growth and Eclat. The associated patterns with lift values greater than 1 indicate the antecedent occurrence had a positive relationship to the consequent occurrence. Thus, this study considered the associated pattern with a lift value greater than 1. Then, the associated patterns obtained from both FP-Growth and Eclat algorithms are considered to see if any of the patterns are the same. Also, only the same patterns from both algorithms were considered, called strong patterns, because they were the most common patterns in the dataset derived from both algorithms.

### *3-3- Identifying Associated Future Earnings Patterns*

We considered only associated future earnings patterns from the strongly associated patterns because of increased and decreased earnings. The associated future earnings patterns can also be used to find the factors that lead to increased or decreased earnings.

## 4- Experimental Setup and Results

### *4-1- Data Description*

A total of 2850 financial statements of 570 companies listed on the Stock Exchange of Thailand from 2016 to 2020 were computed as 68 accounting descriptor variables given in Table 1 and their earnings in the subsequent fiscal year from Equation 1. Data for all five fiscal years was required to cover the calculations according to Equation 1. Table 2 presents examples of descriptor variables and the earnings results from two companies.

**Table 2. Example of descriptor variables and earnings results from two companies**

| Accounting descriptors and earnings result | Company 1 | Company 2 |
|---|---|---|
| 1. Current ratio | 1.5522 | 2.8611 |
| 2. %Δ in current ratio | (24.7311) | 22.8654 |
| 3. Quick ratio | 0.4197 | 1.3684 |
| 4. %Δ in quick ratio | 157.1377 | 35.7223 |
| 5. Days sales in accounts receivable | 27.2192 | 65.3587 |
| 6. %Δ in days sales in accounts receivable | 22.2352 | (24.4299) |
| 7. Inventory turnover | 178.4963 | 199.5043 |
| 8. %Δ in inventory turnover | (22.7713) | (8.1212) |
| 9. Inventory/total assets | 0.3630 | 0.2968 |
| 10. %Δ in inventory/total assets | (48.4076) | (13.6749) |
| 11. %Δ in inventory | (3.9575) | (20.2344) |
| 12. %Δ in sales | 28.6471 | (6.4668) |
| 13. %Δ in depreciation | (3.8464) | 4.0725 |
| 14. Δ in dividend per share | (1.0000) | 0.1333 |
| 15. Depreciation/plant assets | 0.0353 | 0.1961 |
| 16. %Δ in depreciation/plant assets | (78.8767) | 17.2180 |
| 17. Return on opening equity | (0.1144) | 0.0609 |
| 18. Δ in return on opening equity | (0.0576) | 0.0095 |
| 19. %Δ in (capital expenditure/total assets) | 208.5152 | 9.2801 |
| 20. %Δ in (capital expenditure/total assets), one-year lag | 84.7717 | (25.2491) |
| 21. Debt-equity ratio | 0.9693 | 0.4366 |
| 22. %Δ in debt-equity ratio | 42.9004 | (20.8981) |
| 23. Long term debt to equity | 0.1228 | 0.1021 |
| 24. %Δ in long term debt to equity | 189.1480 | 5.4079 |
| 25. Equity to fixed assets | 1.6853 | 4.1277 |

| Accounting descriptors and earnings result | Company 1 | Company 2 |
|---|---|---|
| 26. %Δ in equity to fixed assets | (73.8229) | 12.4282 |
| 27. Times interest earned | (4.3450) | 22.0836 |
| 28. %Δ in times interest earned | (29.5791) | (2.4788) |
| 29. Sales/total assets | 0.8074 | 0.8082 |
| 30. %Δ in sales/total assets | (30.8929) | 1.2249 |
| 31. Return on total assets | (0.0366) | 0.0425 |
| 32. Return on closing equity | (0.0960) | 0.0610 |
| 33. Gross margin ratio | 0.0618 | 0.2429 |
| 34. %Δ in gross margin ratio | 6.2311 | (11.6778) |
| 35. Operating profit (before depreciation) to sales | (0.0441) | 0.1052 |
| 36. %Δ in operating profit (before depreciation) to sales | (84.1883) | 4.5072 |
| 37. Pretax income to sales | (0.0367) | 0.0408 |
| 38. %Δ in pretax income to sales | (37.5994) | 92.4289 |
| 39. Net profit margin | (0.0454) | 0.0525 |
| 40. %Δ in net profit margin | (42.0518) | 23.4993 |
| 41. Sales to total cash | 20.8463 | 3.9886 |
| 42. Sales to accounts receivable | 8.8638 | 6.9697 |
| 43. Sales to inventory | 2.2245 | 2.7229 |
| 44. %Δ in sales to inventory | 33.9481 | 17.2602 |
| 45. Sales to working capital | 4.5289 | 1.8653 |
| 46. %Δ in sales to working capital | 55.9898 | (8.9895) |
| 47. Sales to fixed assets | 3.5678 | 4.7923 |
| 48. %Δ in production | 28.1518 | (2.3243) |
| 49. %Δ in R&D | 78.1141 | (4.5561) |
| 50. %Δ in (R&D/sales) | 0.0038 | 0.0002 |
| 51. %Δ in advertising expense | 213.0590 | (13.0667) |
| 52. %Δ in (advertising/sales) | 0.0143 | (0.0007) |
| 53. %Δ in total assets | 0.8616 | (0.0760) |
| 54. Cash flow to total debt | 0.1128 | 0.3432 |
| 55. Working capital/total assets | 0.1782 | 0.4333 |
| 56. %Δ in working capital/total assets | (55.6977) | 11.2234 |
| 57. Operating income/total assets | (0.0356) | 0.0850 |
| 58. %Δ in operating income/total assets | (27.2872) | 5.7874 |
| 59. %Δ in total uses of funds | (50.0000) | 93.6699 |
| 60. %Δ in total sources of funds | 6,828.9168 | (98.8541) |
| 61. Repayment of long term debt as % of total long term debt | (0.1008) | (0.0026) |
| 62. Issuance of long term debt as % of total long term debt | 0.3382 | 0.0156 |
| 63. Purchase of treasury stock as % of stock | 2.3872 | 0.4774 |
| 64. %Δ in funds | (0.0164) | (0.2336) |
| 65. %Δ in long term debt | 0.3427 | (0.0080) |
| 66. Cash dividend as % of cash flows | (1.6858) | (36.5686) |
| 67. %Δ in working capital | 1.3535 | 0.0000 |
| 68. Net income over cash flows | (220,805.0300) | (206,712.0000) |
| **Earning Result** | **Increase** | **Decrease** |

### *4-2- Data Optimized Discretization*

The dataset was discretized using the Python Optimal Binning (OptBinning) library [26]. The method is used to find optimal bins for each descriptor. Tables 3 and 4 show an example of binning information of "current ratio" and "debt-to-equity ratio" descriptors, respectively. The summary of optimal bins of all descriptors was presented in Table 5.

**Table 3. Example of optimal bins of "current ratio" descriptors**

| Bin | Current ratio range | Count (%) |
|---|---|---|
| 0 | (0.00, 0.66) | 13.1336 |
| 1 | [0.66, 1.71) | 38.7097 |
| 2 | [1.71, 1.89) | 5.5300 |
| 3 | [1.89, 2.34) | 7.1429 |
| 4 | [2.34, 3.60) | 13.5945 |
| 5 | [3.60, 8.83) | 16.1290 |
| 6 | [8.83, inf) | 5.7604 |

**Table 4. Example of optimal bins of "debt-equity ratio" descriptors**

| | Debt-equity ratio range | Count (%) |
|---|---|---|
| 0 | (0.00, 0.11) | 6.4516 |
| 1 | [0.11, 1.22) | 68.8940 |
| 2 | [1.22, 1.54) | 7.8341 |
| 3 | [1.54, 2.23) | 10.8295 |
| 4 | [2.23, inf) | 5.9908 |

**Table 5. Summary of optimal bins of all descriptors**

| No. | Accounting descriptors | Number of bins |
|---|---|---|
| 1 | Current ratio | 7 |
| 2 | %Δ in current ratio | 7 |
| 3 | Quick ratio | 8 |
| 4 | %Δ in quick ratio | 8 |
| 5 | Days sales in accounts receivable | 7 |
| 6 | %Δ in days sales in accounts receivable | 7 |
| 7 | Inventory turnover | 5 |
| 8 | %Δ in inventory turnover | 7 |
| 9 | Inventory/total assets | 7 |
| 10 | %Δ in inventory/total assets | 4 |
| 11 | %Δ in inventory | 7 |
| 12 | %Δ in sales | 7 |
| 13 | %Δ in depreciation | 8 |
| 14 | Δ in dividend per share | 5 |
| 15 | Depreciation/plant assets | 5 |
| 16 | %Δ in depreciation/plant assets | 4 |
| 17 | Return on opening equity | 7 |
| 18 | Δ in return on opening equity | 8 |
| 19 | %Δ in (capital expenditure/total assets) | 7 |
| 20 | %Δ in (capital expenditure/total assets), one-year lag | 7 |
| 21 | Debt-equity ratio | 5 |
| 22 | %Δ in debt-equity ratio | 4 |
| 23 | Long term debt to equity | 7 |
| 24 | %Δ in long term debt to equity | 7 |
| 25 | Equity to fixed assets | 8 |
| 26 | %Δ in equity to fixed assets | 5 |
| 27 | Times interest earned | 6 |
| 28 | %Δ in times interest earned | 10 |
| 29 | Sales/total assets | 6 |
| 30 | %Δ in sales/total assets | 8 |
| 31 | Return on total assets | 9 |
| 32 | Return on closing equity | 5 |
| 33 | Gross margin ratio | 7 |

| No. | Accounting descriptors | Number of bins |
|---|---|---|
| 34 | %Δ in gross margin ratio | 8 |
| 35 | Operating profit (before depreciation) to sales | 9 |
| 36 | %Δ in operating profit (before depreciation) to sales | 11 |
| 37 | Pretax income to sales | 10 |
| 38 | %Δ in pretax income to sales | 8 |
| 39 | Net profit margin | 8 |
| 40 | %Δ in net profit margin | 9 |
| 41 | Sales to total cash | 8 |
| 42 | Sales to accounts receivable | 6 |
| 43 | Sales to inventory | 4 |
| 44 | %Δ in sales to inventory | 7 |
| 45 | Sales to working capital | 6 |
| 46 | %Δ in sales to working capital | 7 |
| 47 | Sales to fixed assets | 7 |
| 48 | %Δ in production | 8 |
| 49 | %Δ in R&D | 6 |
| 50 | %Δ in (R&D/sales) | 8 |
| 51 | %Δ in advertising expense | 6 |
| 52 | %Δ in (advertising/sales) | 8 |
| 53 | %Δ in total assets | 8 |
| 54 | Cash flow to total debt | 8 |
| 55 | Working capital/total assets | 8 |
| 56 | %Δ in working capital/total assets | 6 |
| 57 | Operating income/total assets | 10 |
| 58 | %Δ in operating income/total assets | 5 |
| 59 | %Δ in total uses of funds | 5 |
| 60 | %Δ in total sources of funds | 5 |
| 61 | Repayment of long term debt as % of total long term debt | 4 |
| 62 | Issuance of long term debt as % of total long term debt | 5 |
| 63 | Purchase of treasury stock as % of stock | 1 |
| 64 | %Δ in funds | 4 |
| 65 | %Δ in long term debt | 5 |
| 66 | Cash dividend as % of cash flows | 7 |
| 67 | %Δ in working capital | 2 |
| 68 | Net income over cash flows | 5 |

From Table 3, the results show that there are seven optimal ranges of the current ratio descriptor. For example, it can be found from Table 2 that if company 1 has a current ratio = 1.5522, it is discretized into the range [0.66, 1.71), and if company 2 has the current ratio = 2.8611, it is discretized into the range [2.34, 3.60).

From Table 4, the results show five optimal ranges of debt-to-equity ratio descriptor. For example, it can be found from Table 2 that if company 1 has debt-to-equity ratio = 0.9693, it is discretized into the range [0.11, 1.22), and if company 2 has debt-to-equity ratio = 0.4366, it is also discretized into the range [0.11, 1.22).

From Table 5 shows the optimal range of each descriptor using optimized discretization. For example, in Tables 3 and 4, current and debt-to-equity ratios have seven and five optimal ranges, respectively.

### *4-3- Identifying Associated Future Earnings Patterns*

From 2850 financial statements, strongly associated patterns were found using FP-Growth and Eclat algorithms in a Mlxtend library (Machine learning extensions) [27]. The proposed method was applied several times to determine the best parameters for strongly associated patterns. Therefore, the minimum support value and minimum confidence thresholds of the FP-Growth and Eclat algorithms were set to be 15% and 60%, respectively. We found a total of 39 patterns with 90%-100% confidence and a lift value >1 through the FP-Growth algorithm and 34 patterns with 90%-100% confidence and a lift value > 1 from the Eclat algorithm, as shown in Tables 6 and 7. In total, four strongly associated increased earning patterns, and nine strongly associated decreased earning patterns were mined from both algorithms. The strong patterns of associated earnings were shown in Table 8.

**Table 6. Associated patterns of earnings derived from FP-Growth**

| No. | Antecedent | Consequent | Confidence (%) |
|---|---|---|---|
| 1 | %Δ in total uses of funds = (-inf, -740.14) | Earnings = increase | 99.47 |
| 2 | Inventory turnover = (1.00, 87.14) | Earnings = increase | 94.16 |
| 3 | Issuance of long term debt as % of total long term debt = (0.00, 0.08) | Earnings = increase | 91.72 |
| 4 | Debt-equity ratio = [0.11, 1.22) | Earnings = increase | 90.09 |
| 5 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) and return on closing equity = [0.05, inf) and return on opening equity = [0.09, inf) | Earnings = decrease | 100.00 |
| 6 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) and return on opening equity = [0.09, inf) | Earnings = decrease | 100.00 |
| 7 | %Δ in operating income/total assets = (-inf, 18.06) and inventory turnover = (1.00, 87.14) and return on opening equity = [0.09, inf) | Earnings = decrease | 100.00 |
| 8 | %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and repayment of long term debt as % of total long term debt = [-0.06, inf), %Δ in working capital/total assets = [-50.26, 13.59) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 100.00 |
| 9 | %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and %Δ in current ratio = [-9.79, 12.97) | Earnings = decrease | 100.00 |
| 10 | %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and %Δ in equity to fixed assets = [-16.90, 9.07) | Earnings = decrease | 100.00 |
| 11 | %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 100.00 |
| 12 | %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and inventory turnover = (1.00, 87.14) and return on opening equity = [0.09, inf) | Earnings = decrease | 100.00 |
| 13 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and return on opening equity = [0.09, inf) | Earnings = decrease | 100.00 |
| 14 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and return on opening equity = [0.09, inf) | Earnings = decrease | 100.00 |
| 15 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) and return on closing equity = [0.05, inf) | Earnings = decrease | 99.04 |
| 16 | %Δ in operating income/total assets = (-inf, 18.06) and repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in working capital/total assets = [-50.26, 13.59), long term debt to equity = (0.00, 0.11) | Earnings = decrease | 98.65 |
| 17 | %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and net profit margin = [0.05, 0.11) | Earnings = decrease | 98.61 |
| 18 | %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in depreciation/plant assets = [-55.5603, 7.01) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 98.59 |
| 19 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 97.67 |
| 20 | %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in long term debt = [-0.01, 0.01) | Earnings = decrease | 97.30 |
| 21 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in debt-equity ratio = (-inf, 3.17) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 97.18 |
| 22 | %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in debt-equity ratio = (-inf, 3.17) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 97.10 |
| 23 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) and %Δ in (capital expenditure/total assets), one-year lag = [-7.29, 112.36) | Earnings = decrease | 97.10 |
| 24 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) and repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 97.06 |
| 25 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in equity to fixed assets = [-16.90, 9.07) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 97.06 |
| 26 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08), return on closing equity = [0.05, inf) | Earnings = decrease | 96.30 |
| 27 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 96.10 |
| 28 | %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and return on closing equity = [0.05, inf) | Earnings = decrease | 96.08 |
| 29 | %Δ in operating income/total assets = (-inf, 18.06) and repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 96.08 |
| 30 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) and %Δ in funds = [-0.09, 0.03) | Earnings = decrease | 95.95 |
| 31 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 95.92 |

| No. | Antecedent | Consequent | Confidence (%) |
|---|---|---|---|
| 32 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in Working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 95.79 |
| 33 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in net profit margin = [-65.19, -21.93) | Earnings = decrease | 95.65 |
| 34 | %Δ in operating income/total assets = (-inf, 18.06) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in depreciation/plant assets = [-55.5603, 7.01) and %Δ in debt-equity ratio = (-inf, 3.17) | Earnings = decrease | 95.65 |
| 35 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 95.60 |
| 36 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in long term debt = [-0.01, 0.01) | Earnings = decrease | 95.50 |
| 37 | %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and repayment of long term debt as % of total long term debt = [-0.06, inf) | Earnings = decrease | 95.45 |
| 38 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 95.45 |
| 39 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in equity to fixed assets = [-16.90, 9.07) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 95.06 |

**Table 7. Earnings associated patterns derived from Eclat**

| No. | Antecedent | Consequent | Confidence (%) | Lift |
|---|---|---|---|---|
| 1 | %Δ in total uses of funds = (-inf, -740.14) | Earnings = increase | 99.47 | 1.30 |
| 2 | Inventory turnover = (1.00, 87.14) | Earnings = increase | 94.16 | 1.23 |
| 3 | Issuance of long term debt as % of total long term debt = (0.00, 0.08) | Earnings = increase | 91.72 | 1.20 |
| 4 | Debt-equity ratio = [0.11, 1.22) | Earnings = increase | 90.09 | 1.18 |
| 5 | %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and %Δ in depreciation/plant assets = [-55.5603, 7.01) | Earnings = decrease | 99.04 | 1.38 |
| 6 | Issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) | Earnings = decrease | 96.08 | 1.34 |
| 7 | Repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 96.08 | 1.34 |
| 8 | %Δ in operating income/total assets = (-inf, 18.06) and long term debt to equity = (0.00, 0.11) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 95.92 | 1.33 |
| 9 | Debt-equity ratio = [0.11, 1.22) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 95.79 | 1.33 |
| 10 | Repayment of long term debt as % of total long term debt = [-0.06, inf) and debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 95.60 | 1.33 |
| 11 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in long term debt = [-0.01, 0.01) | Earnings = decrease | 95.50 | 1.33 |
| 12 | Issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) | Earnings = decrease | 94.83 | 1.32 |
| 13 | Debt-equity ratio = [0.11, 1.22) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) | Earnings = decrease | 94.74 | 1.32 |
| 14 | Repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) | Earnings = decrease | 94.06 | 1.31 |
| 15 | Issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in debt-equity ratio = (-inf, 3.17) | Earnings = decrease | 93.69 | 1.30 |
| 16 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) | Earnings = decrease | 93.38 | 1.30 |
| 17 | Debt-equity ratio = [0.11, 1.22) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 93.27 | 1.30 |
| 18 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 92.90 | 1.29 |
| 19 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in funds = [-0.09, 0.03) | Earnings = decrease | 92.66 | 1.29 |
| 20 | Repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in equity to fixed assets = [-16.90, 9.07) and %Δ in operating income/total assets = (-inf, 18.06) | Earnings = decrease | 92.63 | 1.29 |
| 21 | Debt-equity ratio = [0.11, 1.22) and depreciation/plant assets = [0.09, 0.15) and %Δ in operating income/total assets = (-inf, 18.06) | Earnings = decrease | 92.55 | 1.29 |
| 22 | %Δ in equity to fixed assets = [-16.90, 9.07) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease | 92.45 | 1.29 |

| No. | Antecedent | Consequent | Confidence (%) | Lift |
|---|---|---|---|---|
| 23 | Debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 92.25 | 1.28 |
| 24 | Issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 92.24 | 1.28 |
| 25 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in funds = [-0.09, 0.03) | Earnings = decrease | 92.20 | 1.28 |
| 26 | %Δ in equity to fixed assets = [-16.90, 9.07) and %Δ in operating income/total assets = (-inf, 18.06) and debt-equity ratio = [0.11, 1.22) | Earnings = decrease | 92.17 | 1.28 |
| 27 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in inventory/total assets = [3.14, inf) | Earnings = decrease | 92.16 | 1.28 |
| 28 | %Δ in current ratio = [-9.79, 12.97) and %Δ in operating income/total assets = (-inf, 18.06) | Earnings = decrease | 91.87 | 1.28 |
| 29 | Repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in operating income/total assets = (-inf, 18.06) and long term debt to equity = (0.00, 0.11) | Earnings = decrease | 91.59 | 1.27 |
| 30 | %Δ in operating income/total assets = (-inf, 18.06) and sales to working capital = [2.84, 13.76) | Earnings = decrease | 91.26 | 1.27 |
| 31 | Sales to accounts receivable = [4.46, 7.85) and %Δ in operating income/total assets = (-inf, 18.06) | Earnings = decrease | 90.91 | 1.26 |
| 32 | Sales/total assets = [0.90, 1.70) and return on opening equity = [0.09, inf) | Earnings = decrease | 90.82 | 1.26 |
| 33 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in return on opening equity = [-0.02, 0.02) | Earnings = decrease | 90.29 | 1.26 |
| 34 | Sales to inventory = [6.82, inf) and return on opening equity = [0.09, inf) | Earnings = decrease | 90.00 | 1.25 |

**Table 8. Strong associated patterns of increased and decreased earnings from both algorithms**

| No. | Antecedent | Consequent |
|---|---|---|
| 1 | %Δ in total uses of funds = (-inf, -740.14) | Earnings = increase |
| 2 | Inventory turnover = (1.00, 87.14) | Earnings = increase |
| 3 | Issuance of long term debt as % of total long term debt = (0.00, 0.08) | Earnings = increase |
| 4 | Debt-equity ratio = [0.11, 1.22) | Earnings = increase |
| 5 | %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) and %Δ in depreciation/plant assets = [-55.5603, 7.01) | Earnings = decrease |
| 6 | Issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and return on closing equity = [0.05, inf) | Earnings = decrease |
| 7 | Repayment of long term debt as % of total long term debt = [-0.06, inf) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease |
| 8 | %Δ in operating income/total assets = (-inf, 18.06) and long term debt to equity = (0.00, 0.11) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease |
| 9 | Debt-equity ratio = [0.11, 1.22) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease |
| 10 | Repayment of long term debt as % of total long term debt = [-0.06, inf) and debt-equity ratio = [0.11, 1.22) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in working capital/total assets = [-50.26, 13.59) | Earnings = decrease |
| 11 | %Δ in operating income/total assets = (-inf, 18.06) and %Δ in long term debt = [-0.01, 0.01) | Earnings = decrease |
| 12 | Debt-equity ratio = [0.11, 1.22) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and %Δ in depreciation/plant assets = [-55.5603, 7.01) | Earnings = decrease |
| 13 | Debt-equity ratio = [0.11, 1.22) and issuance of long term debt as % of total long term debt = (0.00, 0.08) and %Δ in operating income/total assets = (-inf, 18.06) and long term debt to equity = (0.00, 0.11) | Earnings = decrease |

## 5- Discussion

Proposed solutions from the research questions and discussion of associated future earnings patterns.

***RQ1:*** *What machine learning techniques can be used to identify future earnings patterns in financial statements?*

Before identifying future earning patterns, this study used an optimized discretization technique to find optimal bins and generate discrete data in each accounting descriptor variable. Then, both FP-Growth and Eclat algorithms were used to identify the associated future earnings patterns. The strongly associated future earnings patterns were identified concerning the positive or negative changes in earnings. This is a new way for financial statement users to consider earnings that differ from previous studies that only have found factors affecting earnings.

Using FP-Growth and Eclat algorithms to determine whether the trend of earnings is increasing or decreasing, this study found four associated patterns that were the same for both methods, as shown in Table 8. This indicates that the increased earnings pattern can be found in a few limited patterns. This is in line with the conservative principle [28], in which the probability of accounting transactions and revenue resulting in future economic benefits for an entity should be considered with caution. Those benefits can be measured reliably [29]. All four associated increased earnings patterns resulting from our work were listed under the asset and liability category.

In order to identify decreased earnings patterns using the FP-Growth and Eclat algorithms, 35 and 30 associated patterns were obtained, respectively, with the associated patterns of liability, revenue, and expense categories having the most relevance. For example, if %Δ in operating income/total assets = (-inf, 18.06), the issuance of long-term debt as percentage of total long-term debt = (0.00, 0.08), the repayment of long-term debt as percentage of total long-term debt = [-0.06, inf), %Δ in working capital/total assets = [-50.26, 13.59) and long-term debt to equity = (0.00, 0.11), then earnings decrease. It can be seen that the associated decreased earnings patterns reflect information about liability since liability is the entity's current obligation to transfer economic resources as a result of past events [29]. It also resulted in decreased earnings as the company's debt resulted in the incurred debt [30], causing the firm to be obligated to pay the interest expense and debt in the future. It will affect the future performance and liquidity of a business. Besides, the found items were classified as revenue and expense, resulting in net profit or a net loss. The related items result in a decrease in earnings due to net profit or net loss from past to present [5]. These reflect past performance in which the entity operates profitably enough to pay back its owners or creditors, but they also use historical performance data to predict future earnings [5].

***RQ2:*** *What associated patterns and accounting descriptors are related to future increased and decreased earnings after finding future earnings patterns?*

As for research question 1, this study discovered the strongly associated patterns in both algorithms, and there were four and nine strongly associated increased and decreased earnings patterns, respectively, as given in Table 8.

Four strongly associated increased earnings patterns include:

- If %Δ in total uses of funds = (-inf, -740.14), then increased earnings indicate that the entity's use of funds has decreased. This may be due to the decreased acquisition of fixed or other assets, the redemption of long-term debt, or the payment of creditors. These results affect a decrease in amortization or depreciation expense arising from the acquisition of assets. The decrease in expenses results in an earnings increase, consistent with [7], who found that investing cash flow affects predicted future earnings.
- If inventory turnover = (1.00, 87.14), then increased earnings indicate that within one year, the entity's sold inventories are within a high range, meaning that inventories have been turned over many times. This makes inventories have a low risk of becoming spoiled, damaged, or obsolescent [31]. As a result, a business may not have to recognize the expense of spoilage, damage, or obsolescence of a product [32], which can affect the increased earnings of a business.
- If the issuance of long-term debt as a percentage of total long-term debt = (0.00, 0.08), then earnings increase. This pattern shows that raising money to be used as long-term debt is long-term funding. It is long-term fundraising within the accounting period close to the total long-term debt. As a result, a business will reduce the burden of interest expenses incurred in the future. This is also to manage long-term debt to some extent to pay the interest expense and principal in the future, which increases a business' earnings [33].
- If the debt-to-equity ratio = [0.11, 1.22), an earnings increase indicates that the debt-to-equity ratio reflects the entity's financing. This is related to Ooi (2000) [34], who found that the higher the equity capital level of a firm, the lower its debt agency cost. This pattern shows that the entity uses its funds from external sources similar to that of its internal sources, which commits to paying less interest. As a result, the recognition of expenses of a business is lower, increasing its earnings.

For the nine strongly associated decreased earnings patterns, there are accounting descriptors in all patterns. %Δ in operating income/total assets = (-inf, 18.06) is consistent with [19], who showed that total assets could generate an increase or decrease in operating income from the previous year. If the ratio result is a negative value, the total assets can generate a decrease in operating income from the previous year. A positive value of the ratio result means that total assets can generate an increase in operating income from the previous year, which shows that a business can effectively manage the use of total assets. This descriptor can identify future earnings decrease when it co-occurs with the following descriptors:

- If return on closing equity = [0.05, inf) and %Δ in depreciation/plant assets = (-inf, 7.01), then earnings decrease (Pattern 5). The %Δ in operating income/total assets = (-inf, 18.06) co-occurs with the returns to the owner, in the range [0.05, inf), a high percentage of return on ownership. This also coincides with the depreciation rate being in the range (-inf, 7.01), indicating a higher amortization/depreciation expense [35]. Therefore, when the depreciation rate is calculated at a high rate, this causes a business to have more expenses, affecting a future earnings decrease.
- If issuance of long-term debt as percentage of total long-term debt = (-inf, 0.08) and return on closing equity = [0.05, inf), then earnings decrease (Pattern 6). It can be seen that if the ratio of long-term debt increases against total long-term debt in the range (-inf, 0.08), long-term debt increase has a financial cost higher than raising money from the owners. Besides, if it co-occurs with the return the owner receives in the range [0.05, inf), the business profits from using total assets and has a large proportion of return to its owners. This includes incurring more long-

term liabilities. Consequently, companies are obligated to pay interest expenses in the future, resulting in a decrease in earnings due to higher financial costs [36].

- If repayment of long-term debt as a percentage of total long-term debt = [-0.06, inf) and %Δ in working capital/total assets = [-50.26, 13.59), then earnings decrease (Pattern 7), indicating %Δ in operating income/total assets = (-inf, 18.06) coincides with the long-term debt payout ratio compared to total long-term liabilities. The debt payment will result in the company not having enough money to invest in generating more income. Also, this coincides with the decrease in current assets and current liabilities compared to total assets, meaning that a business does not have a different value of current assets and liabilities. This affects the liquidity of the company's operations, resulting in a decrease in earnings [36].
- If the long-term debt to equity = (0.00, 0.11) and %Δ in working capital/total assets = [-50.26, 13.59), then earnings decrease (Pattern 8), showing that when the %Δ in operating income/total assets = (-inf, 18.06) coincides with long-term debt to equity, resulting in long-term debt higher than equity, this means that the entity uses more funds from creditors than owners, causing a business to have higher financial costs. As a result, a business has higher interest expense obligations. When this co-occurs with the %Δ in working capital/total assets with current assets close to current liabilities, which can affect the company's short-term debt payments. As a result, this entity would have higher financial costs and liquidity risk of short-term debt repayment, leading to a decrease in the earnings of a business [36].
- The debt-to-equity ratio = [0.11, 1.22) and the issuance of long-term debt as percentage of total long-term debt = (-inf, 0.08) shows that the %Δ in operating income/total assets = (-inf, 18.06) co-occurs with the leverage to equity ratio in the range of debt close to equity. However, external capital is still used more than internal capital, which has higher financial costs [36]. This also coincides with the issuance of long-term debt compared to total long-term debt, indicating a high financial cost of raising funds compared to the issuance of long-term debt. As a result, a business is burdened with higher interest expenses. Furthermore, earnings decrease when this co-occurs with the %Δ in working capital/total assets = [-50.26, 13.59] (Patterns 9). The value of current assets close to current liabilities is a warning signal about the liquidity of the company's short-term debt repayment. Also, when this co-occurs with the %Δ in depreciation/plant assets = (-inf, 7.01), earnings decrease (Pattern 12). It is a warning sign of increased depreciation expense recognition, resulting in decreased business revenue.
- If repayment of long-term debt as percentage of total long-term debt = [-0.06, inf) and debt-to-equity ratio = [0.11, 1.22) and the %Δ in working capital/total assets = [-50.26, 13.59), then earnings decrease (Pattern 10). This shows that when the %Δ in operating income/total assets = (-inf, 18.06) coincides with long-term debt payments, this is almost the same as total long-term debt. Therefore, this results in a business not having a cash balance between investing more for higher returns. In addition, the debt is close to the owner's equity. Consequently, a business has higher interest costs, including current assets close to the current liability, which affects the liquidity of the short-term debt repayment of a business. When the above patterns co-occur, causing a business to be burdened with increased interest expenses, including the risk of liquidity in a business, resulting in an earnings decrease [36].
- If there is a %Δ in long-term debt [-0.01, 0.01), then earnings decrease (Pattern 11), indicating that %Δ in operating income/total assets = (-inf, 18.06) coincides with the minimal change in long-term debt. Therefore, the entity would have obligations for interest expenses, resulting in an earnings decrease in a business [36].
- If debt-to-equity ratio = [0.11, 1.22) and issuance of long-term debt as percentage of total long-term debt = (-inf, 0.08), and long-term debt to equity = (0.00, 0.11), then earnings decrease (Pattern 13). This indicates that %Δ in operating income/total assets = (-inf, 18.06) co-occurs with the use of more funds from creditors than owners, resulting in higher interest expenses. Also, the issuance of long-term debt close to total debt affects the interest expense of a business in the future, as well as raising funds from long-term debt to the owners. The above patterns result in recognition of higher interest expenses of a business, which affects an earnings decrease in a business [36].

## 6- Conclusions

This work aimed to find associated future earnings patterns from the financial statements of companies listed on the Stock Exchange of Thailand. Sixty-eight accounting descriptors provided the basis to estimate the probability of an increase or decrease in earnings in the subsequent fiscal year. Before generating associated patterns from accounting descriptors, the numerical data were converted into optimal discrete ranges using an optimized discretization technique. The FP-Growth and Eclat algorithms were used to identify strongly associated future earnings patterns. Our work was a novel approach because these associated patterns could identify future earning relationships between accounting descriptors. It differed from other methods, which only predict earnings or identify the factors of future earnings. We found four patterns of strongly associated increased earnings related to the asset and liability categories in financial statements. The other nine strongly associated patterns of decreased earnings were related to the liability, revenue, and expense categories in financial statements. From strongly associated patterns, we also identified ten important accounting

descriptors reflecting future earnings: 1) %Δ in long-term debt, 2) %Δ in debt-to-equity ratio, 3) %Δ in depreciation/plant assets, 4) %Δ in operating income/total assets, 5) %Δ in working capital/total assets, 6) debt-to-equity ratio, 7) issuance of long-term debt as a percentage of total long-term debt, 8) long-term debt to equity, 9) repayment of long-term debt as a percentage of total long-term debt, and 10) return on closing equity. These results are easy to use to provide warning signs for analysis and decision-making in a business and interpret the co-occurrence of future earnings.

This study has the following implications:

- Academic research implications: This work proposes a new model to find the associated increased and decreased earning patterns that differ from previous studies in the literature review.
- Practical implications for internal user decision-making: This newly proposed method can help the internal users plan a business operation strategy in the future more accurately.
- Practical implications for external user decision-making: These associated future earnings patterns and important accounting descriptors are used to help external users provide supporting information for investment decisions, including considering the ability of a business to repay future debt.
- Nevertheless, this study had the following limitations. We used companies listed on the Stock Exchange of Thailand (SET), which was set up as a long-term funding source for large companies. The Market for Alternative Investment (MAI) is a source of funding for small and medium-sized businesses. As these businesses have a market capitalization of different values, this will result in different optimal ranges of accounting descriptors and will also result in different associated increased or decreased earnings patterns.

Future studies will separate datasets into the different business types of listed companies on the Stock Exchange of Thailand so that associated future earnings patterns can be clearly identified for each business type. In addition, other data analytics methods, such as time-series deep learning, should be used for forecasting future earnings.

## 7- Declarations

### *7-1- Author Contributions*

Conceptualization, P.T., and S.S.; methodology, P.T., and S.S.; software, P.T.; validation, P.T., and S.S.; formal analysis, P.T., and S.S.; investigation, P.T., and S.S.; resources, S.S.; data curation, P.T., and S.S.; writing-original draft preparation, P.T., and S.S.; writing-review and editing, P.T., and S.S.; visualization, P.T.; supervision, S.S.; project administration, S.S. All authors have read and agreed to the published version of the manuscript.

### *7-2- Data Availability Statement*

Data sharing is not applicable to this article.

### *7-3- Funding*

The authors received no financial support for the research, authorship, and/or publication of this article.

### *7-4- Ethical Approval*

Permission for the study was obtained from the ethics committee of Walailak University, Thailand (protocol no. WUEC-21-337-01).

### *7-5- Conflicts of Interest*

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

## 8- References

[1] Dechow, P. M. (1994). Accounting earnings and cash flows as measures of firm performance. The role of accounting accruals. Journal of Accounting and Economics, 18(1), 3–42. doi:10.1016/0165-4101(94)90016-7.

[2] Beaver, W. H. (1968). The Information Content of Annual Earnings Announcements. Journal of Accounting Research, 6, 67. doi:10.2307/2490070.

[3] Block, S. B. (1999). A Study of Financial Analysts: Practice and Theory. Financial Analysts Journal, 55(4), 86–92. doi:10.2469/faj.v55.n4.2288.

[4] Greenberg, R. R., Johnson, G. L., & Ramesh, K. (1986). Earnings versus Cash Flow as a Predictor of Future Cash Flow Measures. Journal of Accounting, Auditing & Finance, 1(4), 266–277. doi:10.1177/0148558X8600100402.

[5] Daraghma, Z. (2013). Predicting the Future Accounting Earnings : Empirical Evidence from the Palestine Securities Exchange. Research Journal of Finance and Accounting, 4(17), 193–204.

[6] Arthur, N., Cheng, M., & Czernkowski, R. (2010). Cash flow disaggregation and the prediction of future earnings. Accounting and Finance, 50(1), 1–30. doi:10.1111/j.1467-629X.2009.00316.x.

[7] Shubita, M. F. (2021). The ability of cash flows to predict the earning: Evidence from Jordan. Investment Management and Financial Innovations, 18(4), 36–44. doi:10.21511/imfi.18(4).2021.04.

[8] Ball, R., & Watts, R. (1972). Some Time Series Properties of Accounting Income. The Journal of Finance, 27(3), 663–681. doi:10.1111/j.1540-6261.1972.tb00991.x.

[9] Ou, J. A., & Penman, S. H. (1989). Financial statement analysis and the prediction of stock returns. Journal of Accounting and Economics, 11(4), 295–329. doi:10.1016/0165-4101(89)90017-7.

[10] Ball, R. T., & Ghysels, E. (2018). Automated Earnings Forecasts: Beat Analysts or Combine and Conquer? Management Science, 64(10), 4936–4952. doi:10.1287/mnsc.2017.2864.

[11] Hou, K., van Dijk, M. A., & Zhang, Y. (2012). The implied cost of capital: A new approach. Journal of Accounting and Economics, 53(3), 504–526. doi:10.1016/j.jacceco.2011.12.001.

[12] Gerakos, J. J., & Gramacy, R. B. (2012). Regression-Based Earnings Forecasts. SSRN Electronic Journal. doi:10.2139/ssrn.2112137.

[13] Li, K. K., & Mohanram, P. (2014). Evaluating cross-sectional forecasting models for implied cost of capital. Review of Accounting Studies, 19(3), 1152–1185. doi:10.1007/s11142-014-9282-y.

[14] Azevedo, V., Bielstein, P., & Gerhart, M. (2021). Earnings forecasts: the case for combining analysts' estimates with a cross-sectional model. Review of Quantitative Finance and Accounting, 56(2), 545–579. doi:10.1007/s11156-020-00902-z.

[15] Anand, V., Brunner, R., Ikegwu, K., & Sougiannis, T. (2019). Predicting Profitability Using Machine Learning. SSRN Electronic Journal. doi:10.2139/ssrn.3466478.

[16] Hunt, J. O. S., Myers, J. N., & Myers, L. A. (2022). Improving Earnings Predictions and Abnormal Returns with Machine Learning. Accounting Horizons, 36(1), 131–149. doi:10.2308/HORIZONS-19-125.

[17] Nguyen, J. P. (2020). Man vs. Machine: Comparing Machine Learning and Analysts' Predictions for Earnings. Wharton Research Scholars, University of Pennsylvania, Philadelphia, United States.

[18] Chen, X., Cho, Y. H., Dou, Y., & Lev, B. (2022). Predicting Future Earnings Changes Using Machine Learning and Detailed Financial Data. Journal of Accounting Research, 60(2), 467–515. doi:10.1111/1475-679X.12429.

[19] Ishibashi, K., Iwasaki, T., Otomasa, S., & Yada, K. (2016). Model Selection for Financial Statement Analysis: Variable Selection with Data Mining Technique. Procedia Computer Science, 96, 1681–1690. doi:10.1016/j.procs.2016.08.216.

[20] Sawangarreerak, S., & Thanathamathee, P. (2021). Detecting and analyzing fraudulent patterns of financial statement for open innovation using discretization and association rule mining. Journal of Open Innovation: Technology, Market, and Complexity, 7(2), 128. doi:10.3390/joitmc7020128.

[21] Kamsu-Foguem, B., Rigal, F., & Mauget, F. (2013). Mining association rules for the quality improvement of the production process. Expert systems with applications, 40(4), 1034-1045. doi:10.1016/j.eswa.2012.08.039.

[22] Park, J., Cha, Y., Jassmi, H. Al, Han, S., & Hyun, C. T. (2020). Identification of defect generation rules among defects in construction projects using association rule mining. Sustainability (Switzerland), 12(9), 3875. doi:10.3390/su12093875.

[23] Han, J., Pei, J., Yin, Y., & Mao, R. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. Data Mining and Knowledge Discovery, 8(1), 53–87. doi:10.1023/B:DAMI.0000005258.31418.83.

[24] Zhang, X., Tang, Y., Liu, Q., Liu, G., Ning, X., & Chen, J. (2021). A fault analysis method based on association rule mining for distribution terminal unit. Applied Sciences (Switzerland), 11(11), 5221. doi:10.3390/app11115221.

[25] Zaki, M. J. (2000). Scalable algorithms for association mining. IEEE Transactions on Knowledge and Data Engineering, 12(3), 372–390. doi:10.1109/69.846291.

[26] Navas-Palencia, G. (2020). Optimal binning: mathematical programming formulation. arXiv preprint, arXiv:2001.08025. doi:10.48550/arXiv.2001.08025.

[27] Raschka, S. (2018). MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack. Journal of Open Source Software, 3(24), 638. doi:10.21105/joss.00638.

[28] Penman, S. H., & Zhang, X. J. (2002). Accounting conservatism, the quality of earnings, and stock returns. Accounting Review, 77(2), 237–264. doi:10.2308/accr.2002.77.2.237.

[29] TFRS (2022). Federation of Accounting Professions. Thai financial reporting standard, Bangkok, Thailand. Available online: https://acpro-std.tfac.or.th/standard/42 (accessed on June 2022).

[30] Chu, K., & Ohlson, J. A. (2019). Accruals and Forecasting. SSRN Electronic Journal. doi:10.2139/ssrn.3340355.

[31] Warren, C. S., Reeve, J. M., & Duchac, J.E. (2017). Accounting. (27th Ed.). Cengage Learning, Boston, United States.

[32] Nasution, A. A. (2020). Effect of inventory turnover on the level of profitability. IOP Conference Series: Materials Science and Engineering, 725(1), 012137. doi:10.1088/1757-899X/725/1/012137.

[33] Ahmed, R., & Bhuyan, R. (2020). Capital Structure and Firm Performance in Australian Service Sector Firms: A Panel Data Analysis. Journal of Risk and Financial Management, 13(9), 214. doi:10.3390/jrfm13090214.

[34] Ooi, J. T. l. (2000). Managerial opportunism and the capital structure decisions of property companies. Journal of Property Investment & Finance, 18(3), 316–331. doi:10.1108/14635780010338317.

[35] Caylor, M. L., & Whisenant, S. (2019). Depreciation Choice and Future Operating Performance. International Journal of Accounting and Financial Reporting, 9(1), 89. doi:10.5296/ijafr.v9i1.13997.

[36] Abor, J. (2005). The effect of capital structure on profitability: an empirical analysis of listed firms in Ghana. Journal of Risk Finance, 6(5), 438–445. doi:10.1108/15265940510633505.