Financial Anomalies Detection Method Example*

Ilona Veitaitė[1,*,†] and Audrius Lopata[2,†]

[1] Vilnius University, Institute of Social Sciences and Applied Informatics, Muitinės str. 8, Kaunas, LT-44280, Lithuania
[2] Kaunas University of Technology, Studentų str. 50, Kaunas, LT-51368, Lithuania

Abstract
The aim of this paper is to provide continuous results on research in financial data analysis. Financial processes involve complex procedures concerning the recording and analysis of financial data. Many companies encounter difficulties when handling large amounts of financial data for assessing the current state of the company, planning future strategies, and other purposes. This paper proceeds with the analysis and usage of financial data space dimensions using General Ledger information from specific companies in the Netherlands, also introduces a method for identifying financial anomalies.

Keywords
process mining; data dimensions; finance analytics; financial anomalies.

1. Introduction

Van der Aalst introduced the notion of process mining in 2004 [24]. Process mining, a data analytics technology, aims to extract process-related insights, focusing on analyzing historical data from process executions recorded as event logs [1]. Various process mining technologies, tools, and applications exist, offering evidence-based solutions and aiding in process enhancements. Business process mining is a relatively young and rapidly growing research area focused on analyzing business processes using a range of data mining and machine learning methods applied to event data. Positioned as a bridge between process science and data science, process mining is indispensable for ambitious and fast-expanding manufacturing enterprises operating within the framework of Industry 4.0. It represents a new form of Big Data Analytics [11, 16, 27, 28, 29, 30].

Business processes encompass a significant volume of events captured by information systems. This data comprises details such as event ID, activity, timestamp, and the individual responsible, collectively referred to as “Event Logs” [5, 12, 20]. Process mining emerges as a rising field within Management Information Systems and Computer Science, employing model-driven methodologies within data mining techniques to analyze intricate business processes. Its objective is to comprehend the current process state based on observed system behavior by deriving process models [1, 25, 26].

An event is an origin of a case, i.e., a process instance (e.g., transferring money from a bank account), comprising an activity (e.g., logging into the bank's website) occurring within a timestamp (e.g., the duration of website usage from login to logout) by an originator (the individual executing the task) [1, 25, 26, 13, 23]. Process mining involves discovering a model by constructing a Petri net [1, 26, 27, 28] based on observed processes following the collection of all event logs [5, 6, 7, 8]. Conformance checking is then performed by process mining to demonstrate that the observed model aligns with the modeled process [5, 20].

The concept of process mining aims to uncover, monitor, and enhance genuine processes – those that exist in reality, rather than assumed ones – by extracting insights from event logs readily accessible in contemporary information systems. Process mining includes automated process discovery (extracting process models from event logs), conformance checking (monitoring variations

* *IVUS2024: Information Society and University Studies 2024, May 17, Kaunas, Lithuania*
1,* Corresponding author
† These author contributed equally.
ilona.veitaite@knf.vu.lt (I. Veitaitė); audrius.lopata@ktu.lt (A. Lopata)
0000-0001-9046-0788 (I. Veitaitė); 0000-0003-2302-8252 (A. Lopata)

by comparing models with logs), social network and organizational mining, automated creation of simulation models, model expansion, model rectification, case prediction, and history-based recommendations [7, 25, 26, 27, 28].

There can be defined three process mining characteristics:

- Process mining extends beyond control-flow discovery. The generation of process models from event logs creates creativity among both practitioners and scholars. Consequently, control-flow discovery is frequently perceived as the most captivating aspect of process mining. However, process mining transcends control-flow discovery. On one hand, discovery represents just one of the three fundamental forms of process mining (discovery, conformance, and enhancement). On the other hand, its scope isn't confined solely to control-flow; the organizational, case, and time perspectives also hold significant relevance [7, 10, 26].
- Process mining is not just a specific type of data mining. Process mining can be seen as the "missing link" between data mining and traditional model-driven BPM. Most data mining techniques are not process centric at all. Process models potentially exhibiting concurrency are incomparable to simple data mining structures such as decision trees and association rules. Therefore, completely new types of representations and algorithms are needed [7, 10, 26].
- Process mining is not limited to offline analysis. Process mining techniques extract knowledge from historical event data. Although "post mortem" data is used, the results can be applied to running cases. For example, the completion time of a partially handled client order can be predicted using a discovered process model [7, 10, 26].

2. Process Mining and Audit Related Activities

Financial auditing involves evaluating companies and their processes to ensure the accuracy and reliability of associated information. Audits verify whether business operations comply with established limitations, which may be defined by managers, administrators, or other parties, including legal or company policies. Detecting violations of these rules can reveal instances of fraud, anomalies, risks, or inefficiencies. Traditionally, auditors provide reasonable assurance regarding process compliance by assessing the effectiveness of controls. However, advancements in data recording, such as event logs and transaction databases, have facilitated a shift towards Auditing 2.0. This approach utilizes detailed process information and advanced process mining techniques to evaluate all events within a business process in real-time, significantly altering the role of auditors [3, 10, 11, 12, 15, 19, 30].

Given the large amount of research that has been produced on the use of modern data mining technology in the field of accounting, an obvious question is: can this research be presented in a structurally logical and thematically coherent manner? In an attempt to answer this question in the affirmative, there is proposed several organizing frameworks for the applications of data mining in accounting. Main idea of this kind frameworks is to present the extensive research coherently. Such frameworks enhance understanding of complex relationships within literature, providing a structured way to map research within a domain.

The main goal of descriptive data mining is business and data understanding (the what happened), the goal of predictive data mining is using the past to understand the future (the what could happen), and the goal of prescriptive data mining is to achieve the best outcome (the what should happen). Descriptive data mining focuses on understanding past and present data to make informed decisions, employing techniques to categorize, characterize, and visualize information [3, 12, 21, 30].

Data mining uses various techniques from statistics, machine learning, and databases. Neural networks emerge as the most widely used technique, followed by regression, decision trees, support vector machines, and genetic algorithms. Less common techniques include text mining, self-organizing maps, and Bayesian networks [3, 8, 9, 21, 25].

A topical analysis reveals that the majority of data mining applications in accounting focus on assurance and compliance, followed by managerial accounting and financial accounting/accounting information systems. This distribution may reflect the differing needs for advanced analytics across various branches of accounting, potentially driven by auditing failures, regulatory tightening, and the demand for technological support. [3, 6, 13, 14, 25].

Auditing using historic data, where historic data, represented by event logs of completed cases, serves as a valuable resource for offline auditing. This data can be filtered and queried to create a more manageable and relevant event log for further analysis. Additionally, historic data can be used to discover de facto process models and assess conformance to de jure models, aiding in identifying deviations and potential problems.

Alternatively, auditing can be conducted solely based on models without direct reference to event data. This approach involves comparing de facto and de jure models to identify differences and update existing models accordingly. Models can also be diagnosed for anomalies using conventional analysis techniques and merged with process mining results to create comprehensive simulation models for what-if analysis [2, 4].

Auditing using current event data and advanced IT systems, auditors can monitor processes in real-time and intervene before completion. Techniques such as process mining allow for on-the-fly monitoring, deviation detection, outcome prediction, and recommendation of corrective actions. However, this operational support raises questions about the auditor's independence and potential interference with the process execution [2, 4].

Auditors face the challenge of dealing with various management information systems that are rapidly growing in data. Traditional methods are becoming obsolete for auditing those financial statements generated by these automated processes. Process mining has been introduced in different corporate contexts, but are missing in the field of auditing. Mining and reconstruction of financial process models can be done using process mining methods. In order to use process mining, data should be in the form of an event log [2, 4].

For process mining to become as widely adopted in accounting information systems, accounting professionals must to acknowledge that the proven value of process mining in those research areas indicates that it surely would be anomalous if it is not equally impactful in our own.

3. Aspects of Financial Process Mining

The specification of Financial Process Mining tasks (projects) has fundamental differences from traditional Process Mining. The Process Mining technology is aimed to discovery of process model from process related data records named Eventlog. Financial processes refer to the methods and procedures completed by the Office of Finance. They include, but aren't limited to: Accounting, Budgeting, Planning and other categorized under varied titles depending on the finance policies and procedures. Since each finance department function has a list of finance business processes involved, drawing up process maps can bring a clear understanding of the tasks and people involved [16, 17, 18].

Basic concepts of finance process mining are listed below [16, 17, 18]:

- Financial (accounting) object (FO): any name of the file field (data record field), i.e. the column name of the excel table), except for time attributes.
- Source data: A subset of financial data records, each record being a set of financial objects and their meanings or codes.
- Case: a unique finance object sequence compiled from event log entries
- Case ID: any selected finance object or combination of few finance objects from the financial data record
- Activity ID: any selected finance object or combination of few finance objects from the financial data record, except included to Case ID.
- Event: one financial data record consisting of the following fields: required field with time parameter value (time stamp) and all others called financial objects (with specified value or code)
- Outcome of finance PM: process model of the behavior of a financial object and its and its differences in different time periods, and statistics (key performance indicators).
- Current problem: to reveal the behavior of financial objects in time, according to data clusters (financial statement types, source document types, ledgers and sub-ledger (journals, etc.)).
- Relevant: behavior of data values and its differences in time periods, according to separate groups of financial data.).

- Process Cube: Process cubes are multidimensional space where the event data is presented and organized using different dimensions. Each cell in the process cube corresponds to a set of events which can be used as an input by any process mining technique. This notion is related to the well-known OLAP (Online Analytical Processing) data cubes, adapting the OLAP paradigm to event data through multidimensional process mining. In this way all operations: slice, dice, drill down, drill up can be implemented [4].

There are many PM tools, their environments are very different, so it is too complicated for a financial specialist to use them directly in formulating data analysis tasks. In [16, 17] were presented a user-friendly approach to PM technology implementation for financial data analysis using a multi-dimensional space of financial data.

Figure 1 presents financial data space (FDS) dimensions and their members, which can be covered with particular data from General Ledger prepared for the analysis according to transformation algorithms [16, 17, 18]:

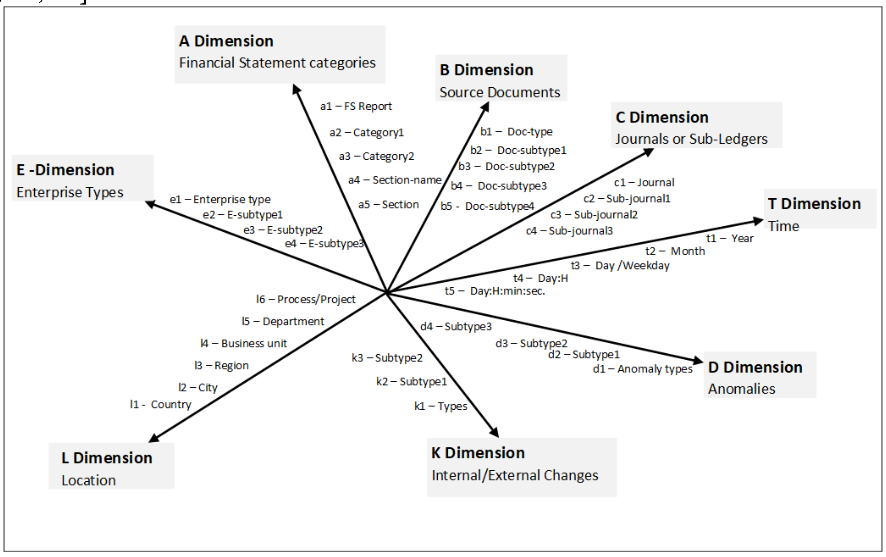


Figure 1: Financial Data Space (FDS) dimensions and dimension members [16, 17, 18]

- A – Dimension – Financial Statement (FS) categories: a1-FS type (Report), a2-CreditCategory1, a3-CreditCategory2, a4-CreditCategory3, a5-SectionCode;
- B – Dimension – Source documents: b1-Doc-Type, b2-Doc-Subtype1, b3-Doc-subtype2, b4-Doc-subtype3, ...;
- C – Dimension – Journals or Sub-Ledgers: c1-Journal, c2-Sub-Journal1, c3-Sub-journal2, c4-Sub-journal3, ...;
- E – Dimension – Enterprise Types: e1-Enterprise Type, e2-E-SubType1, e3-E-SubType2, e4-E-SubType3, ...;
- L – Dimension – Location: l1-Country, l2-City, l3-Region, l4-Business Unit, l5-Department, l6-Process /Project, ...;
- T – Dimension – Time-Period: t1-Year, t2-Month, t3-Day / week day, t4-Day: Hour: min: sec, t5-Hour: min: sec, t6-Period Beginning, t7-Period-Ending;
- D – Dimension – Anomalies: d1-Anomaly type, d2-subtype1, d3-subtype2, d4-subtype3, ...;
- K – Dimension – Changes: Internal / External Internal Changes (IC): k1-types, k2-subtype1, k3-subtype2, ...; External Changes (EC): k1-types, k2-subtype1, k3-subtype2, ...

According to the user's specific need for financial data analysis, the expert selects in the financial data space which FDS dimensions are relevant (will be visible to the PM tool environment) and which dimension members are important for specifying the PM project [16, 17, 18].

Financial Process Cube dimensions may be associated with the different Financial Data Space dimensions in the different way.

Financial Process Cube dimensions for example can be associated with the Financial Data Space dimensions as follows:

- Case type dimension is associated with the Financial Statement Category (dimension A),

- Event class dimension is associated with the Document Type (dimension B),
- Time window dimension is associated with the Financial Period (dimension T).

The Financial Process Mining tool composes the Financial Process Cube according to the user specification and displays (visualizes) PC dimensions and their members, continuing the example:

- Case type dimension is associated with the Financial Statement Category (dimension A) members a1 – FS Category and a5 - Section,
- Event class dimension is associated with the Document Type (dimension B) members, b3 - Doc-subtype3,
- Time window dimension is associated with the Financial Period (dimension T), t3 - FinancialYear.

The Process Cube dimensions and members according to the user requirements specification are as follows: PC = {Case type (a1, a5); (Event class (b3); (Times window (t1)};

The next step is to specify the parameters of the PM project according to the objectives of the analysis performed. We select Case ID, Actitvity ID, and Timestamp ID from existing cube dimensions and their members. The example of PM project specification is as follows:

- Dimension FS Category: Case ID: a1 - Category, a5 - Section Code;
- Dimension Document types: Activity ID: b3 – Doc-subtype3 (Invoice);
- Dimension TimeWindow: Timestamp: t3 – Financial Year.

In this step, according to the project specification, the PM tool creates a project EventLog from the existing data set (i.e. Initial Event), on the basis of which the PM process will be started:

- CaseID =(CaseID1=StatementType AND CaseID2=SectionCode),
- ActivityID = InvoiceNumber (i.e. doc-subtype3),
- Timestamp= FinancialYear.

Case type (CaseID1 and CaseID2,..) can be associated with financial process rules (constraints) defined through data record attributes and their values.

These rules of the financial process make it possible to distinguish between permissible and non-permissible transactions, i.e. allows you to detect inadequate records. The rules of financial processes (constraints) are based on the expert knowledge presented in natural language and then formally specified using expression IF (conditions) THEN (Action) and decision tables.

Constraints for Horizontal dimension (= Activity type) members (ActivityID = doc-subtype3, ...) is based on the expert knowledge (formally specified as decision table or otherwise.

The list of doc-subtype3 possible values: doc-subtype3 = (Invoice, Quote, Order, ...)

Example of the Decision table for ActivityID = doc-subtype3 when Transaction type = (DebitSectionCode – CreditSectionCode).

The PM execution results are the discovered process model, which can be also represented graphically (Process Map), and the process model parameters (static data). With the help of PM tool it is possible to get various statistical information, such as:

- General statistics of data set: Number of records, Cases (number), Variants (number), Max number of Events in CaseID (Longest case), Activity ID of Longest case;
- Case ID, Characteristics (Quantities): (QC1) Duration of case, (QC2) Case Started, (QC3) Case Finished, (QC4) Number of Events in the case, (QC5) A number of Resources for each Activity in the case;
- Activity, Characteristics (Quantities): (QA1) Frequency of Activity (a quantity of the same Activity in the data), Relative quantities (%): (RA1) Relative Frequency of Activity (%);
- Resource, Characteristics (Quantities): (QR1) Median Duration, (QR2) Mean duration, (QR3) Duration Range, (QR4) Cumulative duration, (QR5) Frequency of Resource (a number); Relative quantities (%): (RR1) Relative frequency of Resource (%);
- Attributes, Characteristics (Quantities): (Q01) Frequency, (Q02) Cumulative Frequency; Relative quantities (%): (R01) Relative Frequency of Attribute (%);
- Variant (A Cluster of cases), Characteristics (Quantities): (QV1) Cumulative % of Variants (graphical scheme), (QV2) Median Duration (of cases included in the Variant), (QV3) Mean duration (of cases included in the Variant), (QV4) Number of different Cases in the Variant, (QV5) Number of Events (Activities) in the Variant (in the cluster of cases); Relative

quantities (%): (RV1) Relative frequency (%) of Variant (% of cases with the same sequence of Activities (Events)).

4. Financial Anomalies Detection

Using process mining for anomalies detection in financial data empowers companies to protect their assets, maintain investor confidence, and uphold the integrity and accuracy of financial information.

4.1. Financial Data Set Anomaly Detection

For the companies the process of detecting anomalies in financial data is crucial for identifying potential fraud, errors, or irregularities that could damage the integrity of financial reporting and entire status of the company. Anomalies detection in financial data using process mining techniques provides companies with the ability to proactively mitigate risks and enhance compliance with regulatory requirements. The main steps of financial anomaly detection are listed below [18, 22, 23]:

1. Discovery of a normalized model (company-specific):
   1.1. to detect anomalous journal entries, first to be done is to define "normality" with respect to accounting attribute type, indicator type.
2. Identification of deviations of attribute values:
   2.1. to exhibit unusual or rare individual attribute values. Such anomalies usually relate to skewed attributes, e.g. rarely used ledgers, journals or unusual posting times. Traditionally, "red-flag" tests performed by auditors during an annual audit, are designed to capture this type of anomaly.
3. Unusual or rare combinations of attribute values:
   3.1. journal entries that exhibit an unusual or rare combination of attribute values while their individual attribute values occur quite frequently: e.g. unusual accounting records, irregular combinations of general ledger accounts, user accounts used by several accounting departments.
4. Actual time periods list.

It should be noted that some steps are related to method of detection of anomalies in large scale accounting data where according to the authors, their score accounts for both of the observed characteristics, namely: (1) any "unusual" attribute value occurrence (global anomaly) and (2) any "unusual" attribute value co-occurrence (local anomaly): [an unusual or rare combination of attribute values] [22].

4.2. Example of Anomaly Detection

Before the presentation of financial anomaly detection example some concepts must be defined. Changing behavior of some KPI (key performance indicators) is one of the most important indications of changes in companies' decisions. New type of KPI's for anomaly detection are defined below [23]:

1. BCI-A - Absolute: (change per Financial Period: Financial Year or Month):
   1.1. BCI-A1 – Absolute change against previous Financial Period;
   1.2. BCI-A2 – Absolute change against average of Financial Period;
2. BCI-RO - Robustness Coefficient: measure the distance of KPI from the Normative value;
3. BCI-RE1 – Relative (%): change of KPI per Financial Period, i.e. KPI change compared to previous period;
4. BCI-RE2 – Relative (%): change of KPI per Financial Period compared to the KPI Average value);
5. DELTA of some BCI is defined as a change of BCI of the given period comparing to a change of BCI of the previous period (change of change). DELTA of some BCI indicates the trend (style) of changes.

There is anomaly detection in KPI: Current Ratio (Cr) (KPI which equals current assets / current liabilities) presented in the example. Figure 2 presents fragment of company's financial data.

FinancialYear	FinancialPer	Current assets	Current liabilities	Cr	FinancialYear	FinancialPeric	Current assets	Current liabilities	Cr
2012	1	472926.02	78104.38	6.06	2013	1	772080.06	87746.83	8.80
2012	2	486957.94	73095.36	6.66	2013	2	836866.46	111724.53	7.49
2012	3	519019.91	80509.65	6.45	2013	3	832086.77	87474	9.51
2012	4	524649.39	72745.05	7.21	2013	4	840538.57	78373.84	10.72
2012	5	575144.06	91636.9	6.28	2013	5	858990.82	69933.72	12.28
2012	6	608089.91	79105.12	7.69	2013	6	886035.31	72493.17	12.22
2012	7	599764.93	53196.18	11.27	2013	7	957628.12	64677.8	14.81
2012	8	591843.24	45496.47	13.01	2013	8	969500.24	76070.55	12.74
2012	9	677234.4	68099.03	9.94	2013	9	1038608.25	81929.81	12.68
2012	10	715213.64	73959.95	9.67	2013	10	1131501.34	98990.09	11.43
2012	11	708664.3	54249.06	13.06	2013	11	1198282.48	112406.29	10.66
2012	12	725915.52	50684.45	14.32	2013	12	1254737.06	103644.77	12.11

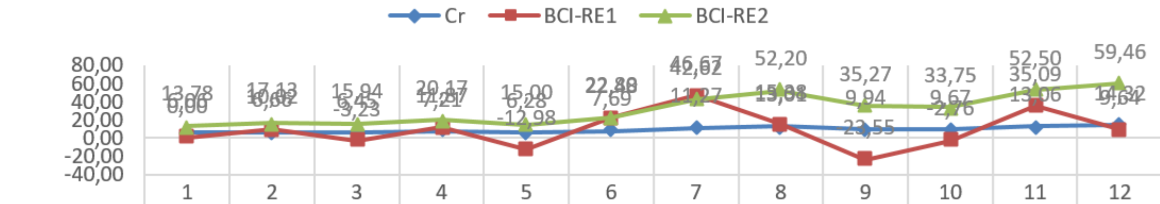
Figure 2: Current Assets, Current Liabilities, Current Ratio

According to BCI described above there are done two variants of Cr calculations.

- BCI-A1 and BCI-RE1 - comparing the current period with the previous period (month);
- BCI-A2 and BCI-RE2 - comparing the current period with the average of the whole period (annual average).

DBID	FinancialYear	FinancialPeri	Current assets	Non-current lia	Current liabilities	Cr	BCI-RE1	BCI-A1	BCI-RE2	BCI-A2
DB1_1017	2012	1	472926.02	128473.98	78104.38	**6.06**	0.00	0.00	13.78	2.49
DB1_1017	2012	2	486957.94	128112.11	73095.36	**6.66**	10.02	0.61	17.13	3.10
DB1_1017	2012	3	519019.91	127378.18	80509.65	**6.45**	-3.23	-0.22	15.94	2.89
DB1_1017	2012	4	524649.39	127380.77	72745.05	**7.21**	11.87	0.77	20.17	3.65
DB1_1017	2012	5	575144.06	127011.24	91636.9	**6.28**	-12.98	-0.94	15.00	2.72
DB1_1017	2012	6	608089.91	127011.24	79105.12	**7.69**	22.48	1.41	22.80	4.13
DB1_1017	2012	7	599764.93	127011.24	53196.18	**11.27**	46.67	3.59	42.62	7.71
DB1_1017	2012	8	591843.24	127011.24	45496.47	13.01	15.38	1.73	52.20	9.45
DB1_1017	2012	9	677234.4	127011.24	68099.03	**9.94**	-23.55	-3.06	35.27	6.38
DB1_1017	2012	10	715213.64	127011.24	73959.95	**9.67**	-2.76	-0.27	33.75	6.11
DB1_1017	2012	11	708664.3	127011.24	54249.06	**13.06**	35.09	3.39	52.50	9.50
DB1_1017	2012	12	725915.52	117239.03	50684.45	**14.32**	9.64	1.26	59.46	10.76

Figure 3: Data Set of each month of 2012

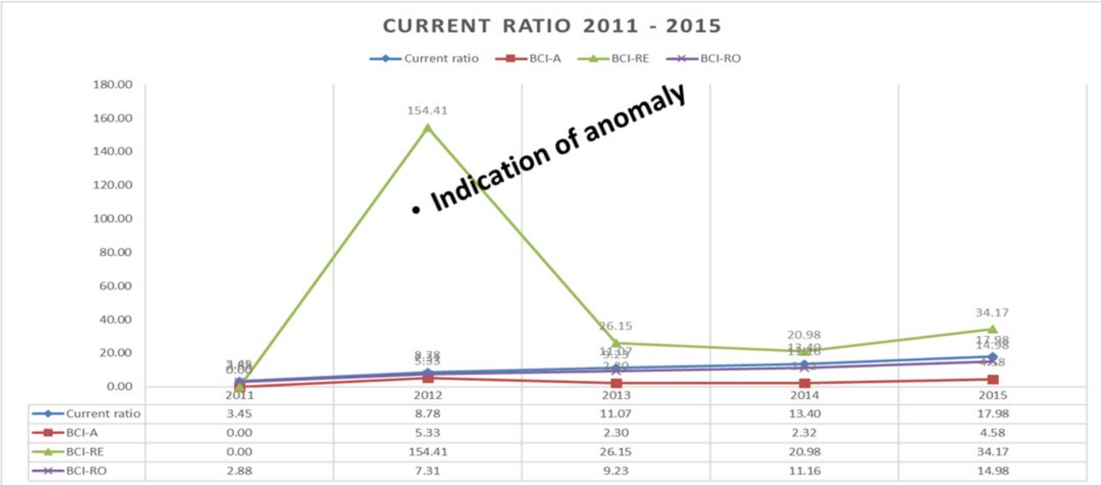


	1	2	3	4	5	6	7	8	9	10	11	12
Cr	6,06	6,66	6,45	7,21	6,28	7,69	11,27	13,01	9,94	9,67	13,06	14,32
BCI-RE1	0,00	10,02	-3,23	11,87	-12,98	22,48	46,67	15,38	-23,55	-2,76	35,09	9,64
BCI-RE2	13,78	17,13	15,94	20,17	15,00	22,80	42,62	52,20	35,27	33,75	52,50	59,46

Figure 4: Current Ratio: BCI-RE1 and BCI-RE2

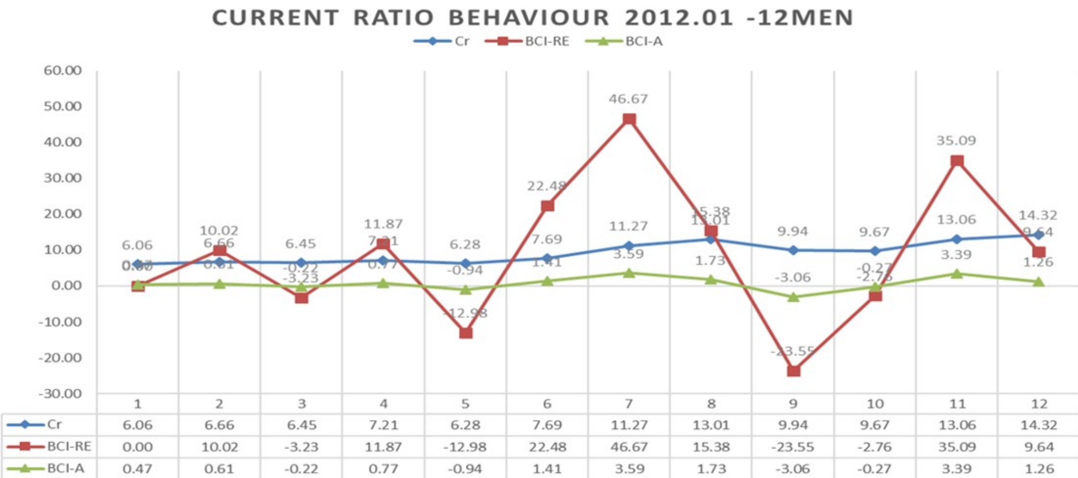
After the calculations (figure 3) and visual diagram (figure 4) anomaly is detected in the seventh financial period of 2012, where BCI-A1 = 3.51, BCI-A2= 7.71 and BCI-RE1 = 46.67%, BCI-RE2 = 42.62%.

Further process of anomaly identification (figure 5) in Current Ration (Cr) behavior using BCI was implemented accordingly: Indications of the Cr behavior anomaly in 2012 were calculated: BCI-RE = 154.41%; Drill down process implemented to data set of 2012: calculation of Cr BCI-A and BCI-RE in the Financial periods 1 – 12 of year 2012; Delta calculation of Cr components values: Current Assets and Current liabilities; and Accounting anomaly scoring: if an entry is anomalous or if it was created by a “regular” business activity.



	2011	2012	2013	2014	2015
Current ratio	3.45	8.78	11.07	13.40	17.98
BCI-A	0.00	5.33	2.30	2.32	4.58
BCI-RE	0.00	154.41	26.15	20.98	34.17
BCI-RO	2.88	7.31	9.23	11.16	14.98

Figure 5: Indication of anomaly in 2012: Cr [average of year] : BCI-RE = 154.41%



	1	2	3	4	5	6	7	8	9	10	11	12
Cr	6.06	6.66	6.45	7.21	6.28	7.69	11.27	13.01	9.94	9.67	13.06	14.32
BCI-RE	0.00	10.02	-3.23	11.87	-12.98	22.48	46.67	15.38	-23.55	-2.76	35.09	9.64
BCI-A	0.47	0.61	-0.22	0.77	-0.94	1.41	3.59	1.73	-3.06	-0.27	3.39	1.26

Figure 6: Cr behavior in 2012: BCI-A and BCI-RE

Drill down process consisted of several elements of Current Ratio (Cr) in each financial period of 2012. The components of Cr are KPIs: Current Assets and Current Liabilities, where Current Ratio (Cr) = Current Assets / Current Liabilities.

Anomaly detection was implemented by identifying it in KPI behavior using BCI (figure 6): firstly, calculation of BCI's of KPI over the period of the year: BCI-A, BCI-RE, BCI-RO; secondly, indication of anomaly of some BCI change in the same year (2011, 2012, … 2015); thirdly, drill down to BCI's of KPI over the Financial periods of the identified year: calculation of BCI-A, BCI-RE and BCI-RO in the financial periods 1 – 12.

5. Conclusions

The paper introduces the benefits of using process mining in financial data analysis, outlining process mining and audit-related tasks while defining aspects of financial process mining. It focuses on the approach of anomaly detection in financial data, using behavioral change indicators (BCIs) based on traditional KPIs. This is part of the results of the project "Platform of tools for the analysis of corporate financial activity data".

The advantages of using BCIs in financial data analysis can be summarized as follows: BCI calculations provide both quantitative and visual insights into KPI behavior over time. Changes in KPI values serve as crucial indicators of changes in company processes or decisions. Using BCIs to track changes in KPIs and identify suspicious trends, reduces the amount of data that needs to be analyzed.

The presented example effectively illustrates the practical implementation of the defined method and underscores its effectiveness in enhancing the understanding and management of financial processes within companies. It provides necessary information for the auditor to make further decisions.

6. Acknowledgements

This paper presents part of results of research project "Enterprise Financial Performance Data Analysis Tools Platform (AIFA)". The research project is funded by European Regional Development Fund according to the 2014–2020 Operational Programme for the European Union Funds' Investments under measure No. 01.2.1-LVPA-T-848 "Smart FDI". Project no.: 01.2.1-LVPA-T-848-02-0004; Period of project implementation: 2020-06-01 - 2022-05-31

7. References

[1] Adriansyah, A. A., & Buijs, J. J. (2013). Mining Process Performance from Event Logs. In *Lecture notes in business information processing* (pp. 217–218). https://doi.org/10.1007/978-3-642-36285-9_23

[2] Alrefai, A., (2019) Audit Focused Process Mining: The Evolution Of Process Mining And Internal Control, PhD Thesis, 2019, https://rucore.libraries.rutgers.edu/rutgers-lib/60514/PDF/1/play/

[3] Amani, F., & Fadlalla, A. (2017). Data mining applications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, 24, 32–58. https://doi.org/10.1016/j.accinf.2016.12.004

[4] Bolt, A. A., & Van Der Aalst, W. M. P. (2015). Multidimensional process mining using process cubes. In *Lecture notes in business information processing* (pp. 102–116). https://doi.org/10.1007/978-3-319-19237-6_77

[5] Brzychczy, E. (2017). Wykorzystanie eksploracji procesów w przedsiębiorstwie. *Inżynieria Mineralna*. https://doi.org/10.29227/im-2017-02-26

[6] Dakić, D., Stefanović, D., Ćosić, I., Lolić, T., & Medojević, M. (2018). Business Process Mining Application: A Literature Review. In *Annals of DAAAM for ... & proceedings of the International DAAAM Symposium* (pp. 0866–0875). https://doi.org/10.2507/29th.daaam.proceedings.125

[7] Daniel, F., Barkaoui, K., & Dustdar, S. (2012). Business Process Management workshops: BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers. Springer Science & Business Media.

[8] Das, K., & Schneider, J. (2007). Detecting anomalous records in categorical datasets. KDD '07: *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/1281192.1281219

[9] Debreceny, R., & Gray, G. L. (2010). Data mining journal entries for fraud detection: An exploratory study. *International Journal of Accounting Information Systems*, 11(3), 157–181. https://doi.org/10.1016/j.accinf.2010.08.001

[10] Earley, C. E. (2015). Data analytics in auditing: Opportunities and challenges. *Business Horizons*, 58(5), 493–500. https://doi.org/10.1016/j.bushor.2015.05.002

[11] Gepp, A., Linnenluecke, M. K., O'Neill, T., & Smith, T. (2017). Big data techniques in auditing research and practice: Current trends and future opportunities. *Social Science Research Network*. https://doi.org/10.2139/ssrn.2930767

[12] Gehrke, N. (n.d.). Basic Principles of Financial Process Mining A Journey through Financial Data in Accounting Information Systems. AIS Electronic Library (AISeL). http://aisel.aisnet.org/amcis2010/289

[13] *How is Process Mining Different From. . . — Flux Capacitor.* (n.d.). https://fluxicon.com/blog/2014/02/how-is-process-mining-different-from/

[14] Jans, M., Alles, M., & Vasarhelyi, M. A. (2013). The case for process mining in auditing: Sources of value added and areas of application. *International Journal of Accounting Information Systems*, 14(1), 1–20. https://doi.org/10.1016/j.accinf.2012.06.015

[15] Janvrin, D. J., & Watson, M. W. (2017). "Big Data": A new twist to accounting. *Journal of Accounting Education*, 38, 3–8. https://doi.org/10.1016/j.jaccedu.2016.12.009

[16] Lopata, A., Butleris, R., Gudas, S., Rudžionis, V., Rudžionienė, K., Žioba, L., Veitaitė, I., Dilijonas, D., Grišius, E., & Zwitserloot, M. (2021). Financial data preprocessing issues. In *Communications in computer and information science* (pp. 60–71). https://doi.org/10.1007/978-3-030-88304-1_5
[17] Lopata, A., Butleris, R., Gudas, S., Rudžionienė, K., Žioba, L., Veitaitė, I., Dilijonas, D., Grišius, E., & Zwitserloot, M. (2022). Financial Process mining characteristics. In *Communications in computer and information science* (pp. 209–220). https://doi.org/10.1007/978-3-031-16302-9_16
[18] Lopata, A., Gudas, S., Butleris, R., Rudžionis, V., Žioba, L., Veitaitė, I., Dilijonas, D., Grišius, E., & Zwitserloot, M. (2022). Financial data anomaly discovery using behavioral change indicators. *Electronics*, 11(10), 1598. https://doi.org/10.3390/electronics11101598
[19] Moffitt, K. C., & Vasarhelyi, M. A. (2013). AIS in an age of big data. *Journal of Information Systems*, 27(2), 1–19. https://doi.org/10.2308/isys-10372
[20] *Realizing a process cube allowing for the comparison of event data.* (n.d.). Eindhoven University of Technology Research Portal. https://research.tue.nl/en/studentTheses/realizing-a-process-cube-allowing-for-the-comparison-of-event-dat
[21] Roubtsova, E., & Wiersma, N. (2018). A Practical Extension of Frameworks for Auditing with Process Mining. *ENASE 2018 : 13th International Conference on Evaluation of Novel Approaches to Software Engineering*. https://doi.org/10.5220/0006798904060415
[22] Schreyer, M., Sattarov, T., Borth, D., Dengel, A., Reimer, B. (2017) Detection of Anomalies in Large-Scale Accounting Data using Deep Autoencoder Networks. Cornell University. https://doi.org/10.48550/arXiv.1709.05254
[23] Rudžionis, V., Lopata, A., Gudas, S., Butleris, R., Veitaitė, I., Dilijonas, D., Grišius, E., Zwitserloot, M., & Rudžionienė, K. (2022). Identifying irregular financial operations using accountant comments and natural language processing techniques. *Applied Sciences*, 12(17), 8558. https://doi.org/10.3390/app12178558
[24] Van Der Aalst, W. M. P., & Weijters, A. J. M. M. (2004). Process mining: a research agenda. *Computers in Industry*, 53(3), 231–244. https://doi.org/10.1016/j.compind.2003.10.001
[25] Van Der Aalst, W. M. P., Van Hee, K., Van Der Werf, J. M. E. M., & Verdonk, M. (2010). Auditing 2.0: Using process mining to support tomorrow's auditor. *IEEE Computer*, 43(3), 90–93. https://doi.org/10.1109/mc.2010.61
[26] Van Der Aalst, W. M. P. (2011). Process mining. In *Springer eBooks*. https://doi.org/10.1007/978-3-642-19345-3
[27] Van Der Aalst, W. M. P. (2012). Process mining. *ACM Transactions on Management Information Systems*, 3(2), 1–17. https://doi.org/10.1145/2229156.2229157
[28] *Getting started — Process Mining Book 3.0.* (n.d.). https://fluxicon.com/book/read/gettingstarted/
[29] Van der Aalst, W.M.P. (2013). Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining. In: Song, M., Wynn, M.T., Liu, J. (eds) Asia Pacific Business Process Management. AP-BPM 2013. *Lecture Notes in Business Information Processing*, vol 159. Springer, Cham. https://doi.org/10.1007/978-3-319-02922-1_1
[30] Werner, M. (2017). Financial process mining - Accounting data structure dependent control flow inference. *International Journal of Accounting Information Systems*, 25, 57–80. https://doi.org/10.1016/j.accinf.2017.03.004