Reflections on Methods of Digital Processing of Very Large Amounts of Accounting Information

Traian Cristin Nicolae
"Ovidius" University of Constanța, Faculty of Economic Sciences, Romania
traian.nicolae.profesor@gmail.com

Abstract

The purpose of this paper is to briefly explore the applicability of information technology methods to very large volumes of accounting data. Current developments regarding the digitization of various sectors of the economy and society are converging towards changes in the accounting information systems of entities. The possibilities of multi-dimensional processing of accounting input data open up new options for modeling accounting reports. The development in a multi-dimensional system of financial accounting reports presents unexpected opportunities for enriching the information transmitted to the users of the annual financial statements. The applicability of new information technology methods is marked by: widespread use of Internet-based applications, 5G and soon 6G high-speed data transmission systems, Wi-Fi 6e wireless systems, Wi-Fi 7, artificial intelligence applications.

Key words: accountings paradigms, new technology
J.E.L. classification: M41, F60

1. Introduction

The use of digital technologies allows the processing of an abundance of data sources - a lot of disparate information, seemingly unrelated to each other. There is a growing trend to integrate intelligent processes into applications (algorithms, mathematical models, artificial intelligence), based on high computing capabilities. A specific characteristic of these technologies is the inclusion in the analysis process of unstructured data in different formats. (www.hyperledger.org)

2. Theoretical background regarding the application of information technology methods in accounting

There is a whole series of algorithms, developed by companies, research institutes, universities, etc., which generally fall into the following broad classes (www.hyperledger.org):

Linear Regression - Linear regression uses the relationship between two sets of continuous quantitative measures. The first set is called the predictor or independent variable. The other is the response or dependent variable. The purpose of linear regression is to identify the relationship in the form of a formula that describes the dependent variable in terms of the independent variable. Once this relationship is quantified, the dependent variable can be predicted for any instance of an independent variable. (www.hyperledger.org)

Logistic Regression - Logistic regression looks similar to linear regression, but is focused on classification instead of quantitative predictions. Logistic regression checks whether an instance of an input variable falls into a particular category or not. The logistic regression result falls within the range 0 - 1. A value closer to 1 indicates that the input variable falls more clearly into that category, and one closer to 0 indicates the opposite. (www.hyperledger.org)

Classification and Regression Trees - Classification and regression trees use a data classification decision. Each decision is based on a question related to one of the input variables. With each question and corresponding answer, the point-in-time value of the data is classified in a specific way. This set of questions and answers, as well as subsequent divisions of the data create a tree-like

structure. After all these trees complete their evaluation of the data, the process combines the individual results to create a final prediction of the analyzed category.

K-Nearest Neighbors - is also a classification algorithm, which allows multivariate classification of data. The learning process is composed of the training set of stored data. Algorithms in this class are often chosen because they are easy to use and have easy-to-interpret results. (www.hyperledger.org)

Pricing algorithms are increasingly used tools by companies to automatically set prices for their products. Algorithmic valuation (algorithm-based valuation) is a term that generally refers to automated decision-making in the field of value management using rule-based or self-learning algorithms. Algorithmic values are often based on statistical analysis of transactional, inventory and catalog data, as well as external criteria such as market values of competing entities. In some cases, image recognition technologies are also used to use text and visual data such as product descriptions and images. Algorithms track all competitor prices and may implement, in response, changes to the values of the products traded by the user of the algorithm. (www.hyperledger.org)

Pricing algorithms can be used in different ways to set values. (www.hyperledger.org)

Dynamic pricing - Algorithms enable frequent and accurate price changes, including basic, promotional and special offers.

Personalized pricing - The algorithmic approach helps to segment or personalize them, based on data on differences in tastes, price sensitivities and other statistical information held about potential buyers.

Differentiated pricing strategies. Algorithms help optimize differentiated pricing strategies, by product and category, based on consumer perception of the product, similarity to other products, and other factors.

Classification of articles. Classification, filtering or ranking algorithms create certain lists by selections from databases or sort a number of items according to predefined criteria.

Algorithms used in pricing can be:

Static - sets prices based on known formulas and variables, which do not change over time;

Dynamic, machine learning - programmed to achieve a certain goal, for example, maximizing sales, profit, being able to automatically adapt their data processing parameters, based on past results and as they have access to more information.

Artificial intelligence (AI) is the ability of machines or computers to perform tasks and tasks commonly associated with human intelligence. (www.hyperledger.org)

In its broadest sense, AI includes any technology designed to mimic, in one way or another, the way a human being works. AI technology learns, adapts to the surrounding conditions, by assimilating and processing information and storing it for use in other similar situations. There are a number of technologies used to implement AI algorithms, presented below.

Robotic Process Automation (RPA): technology that extracts the list of rules and actions to be performed, watching the user perform a specific task. (www.hyperledger.org)

Expert Systems: a computer program that imitates the human decision-making process.

Computer Vision (CV): methods for making sense of digital images (usually divided into activity, image, and "machine vision" recognition). (www.hyperledger.org)

Natural Language Processing (NLP): handles natural language data (involves language understanding, language generation and machine translation). (www.hyperledger.org)

Neural Networks (NNs or ANNs): a class of algorithms modeled after the neural structure of the human brain that improve their performance without being explicitly trained how to do so. The two subclasses of NNs are Deep Learning (a multi-layer neural network) and Generative Adversarial Networks (GAN - two networks that train each other).

Autonomous Systems: lies at the intersection of robotics and intelligent systems (e.g., intelligent perception, dexterous object manipulation, plan-based robot control, etc.).

Inductive Logic Programming (ILP): uses formal logic to represent a database of facts and formulate hypotheses derived from that data. (www.hyperledger.org)

Decision Networks: is a generalization of the most well-known Bayesian networks/ inference/ Bayesian networks/ inferences, which represent a set of variables and their probabilistic relationships through a map (also called a directed acyclic graph). www.hyperledger.org)

Probabilistic Programming: a framework that does not force you to modify the source code of the program for any change in the input data, but works with probabilistic models. Bayesian Program Synthesis (BPS) is a form of probabilistic programming, where Bayesian programs write new programs of this kind. (www.hyperledger.org)

3. Research methodology. The evolution of computer methods in the processing of accounting data - empirical aspects of the research

The research methodology used in this article consists of empirically researching a data set available on the Internet and building models of accounting records.

In the case of Big Data (BD) technologies, data storage is provided by an infrastructure specially designed for recording, managing and retrieving massive amounts of files and data objects. A BD platform is built for much greater scale, speed and performance than traditional data storage has allowed for an entity. It must allow large volumes of data to be stored and sorted in such a way that it can be easily accessed, used and processed by specific applications. Storage of large volumes of data must be able to scale flexibly as needed. (www.hyperledger.org)

A misconception about big data is that the term only refers to the size of the data set. While this is true in principle, the science behind big data goes deeper. The intent is to extract specific subsets of data from multiple large storage volumes. This data may be widely dispersed across different systems and may not have an obvious correlation. The goal is to unify the data in an intelligent way to enable rapid analysis. From the point of view of big data storage, the analysis can be carried out in two dimensions closely interconnected and dependent on each other, respectively: hardware infrastructure and software infrastructure.

There are three cloud implementation models, related to the form of ownership, presented below. (www.hyperledger.org)

The public cloud - based on the investments of a large software company and intended for global consumers regardless of size and field of activity. Today, most consumers own smartphones, with much of their data—photos, contacts, messages, backups—stored in the cloud. Apple, Google, and Microsoft are the largest public cloud providers for mobile devices, for storage, and for purchasing applications. At the same time, the companies Google and Microsoft offer free of charge packages of Office-type document creation and processing applications to customers. The customer receives those services ostensibly for free, but the providing firms recoup some of the maintenance costs through contextual advertising. Other times they are considered loyalty costs: a service offered to a person for free makes them go to the same product line: phone type, operating system, service provider and others. (www.hyperledger.org)

Private cloud – based on the investments of a company or a conglomerate of vertically integrated companies, intended mostly exclusively for internal users of the company. In this form of organization, the equipment and licenses belong to the owning company and the maintenance, development and support is outsourced to the specialized company. (www.hyperledger.org)

The hybrid cloud – based on the use of services provided by the public cloud interconnected with internal information entities, it is mostly intended for very large companies and aims to expand certain internal processing capabilities in order to provide services to consumers inside the company. (www.hyperledger.org)

Because cloud computing services can only be provided over the Internet, businesses must have access to it in order to use them. In 2022 all companies in the EU (99%) had access to the Internet. However, significant differences can be observed between countries. Thus, in Finland, Sweden, Denmark, the Netherlands, Ireland, the United Kingdom and Belgium, at least 60% of enterprises used cloud computing in 2022. On the other hand, in Romania and Bulgaria only 15% or fewer enterprises did so work - according to public data. (www.ziare.ro)

Of the businesses that reported using cloud computing in 2018, about 69% relied on a cloud solution for their email services. Instead of creating a server infrastructure for their email system, which would have involved, among other things, capital expenditures and maintenance costs, these firms opted for a cloud solution based on operating costs per user. (www.ziare.ro)

Cloud computing services can meet a wide range of other business information technology needs. In 2022 - according to publicly available information - almost 8 out of 10 enterprises (80%) using the cloud used it to store files in electronic format. About 63% reported using it for office software (eg word processors, spreadsheets, etc.), while 48% used it for database hosting. Furthermore, through the cloud, businesses access relatively more advanced software applications for end customers, for example for finance/accounting and for managing information about their customers (customer relationship management - CRM) (47% and 35% respectively). Additionally, 20% reported using cloud computing platforms (typically high-performance) for computing power to run their own business software applications. (www.ziare.ro)

4. Findings. Synthetic review on the application of computer methods in the accounting information system

According to publicly available data, the implementation of digital transformation programs in the banking sector is a business priority, starting in 2019. Currently, traditional banks face, on the one hand, increased competition in certain segments, due to the entry of new actors on the market – digital/FinTech banks, banking platform providers – and, on the other hand, with the rapid change in customer expectations, directed towards digital banking to the detriment of traditional banking (with branches and physical agencies). (www.bnr.ro)

According to publicly available data, banks' decision to invest in technology was mainly motivated by their objective to strengthen their competitive position and market share, to increase their ability to attract and retain customers and to obtain cost reductions and an increase in operational efficiency. (www.bnr.ro)

In order to identify the degree of implementation of big data technologies in the Romanian banking sector, the information published by commercial banks was analyzed during the study.

Aspects related to the types of applications/technologies used in the current activity of banks, the types of IT solutions, algorithms, data (structured, unstructured) used, data sources (internal, external), the value of investments in recent years were targeted. , carried out for the development of IT solutions/applications, the benefits obtained from the implementation and by the providers/developers of the solutions, the monitoring tools of the social platforms used. (www.bnr.ro)

It was also aimed at detailing the IT technologies used for the following types of activities (www.bnr.ro) :

- risk management (credit risk);
- fraud detection/administration;
- resource planning;
- the decision-making process;
- simplifying work processes;
- real-time monitoring of transactions;
- behavioral/prediction/optimization analytics;
- personalized offers;
- Internet banking and mobile access to banking data by customers;
- identifying customer spending patterns;
- customer segmentation by category;
- reporting/statistics;
- real-time monitoring of social networks;
- security.

From the review of the published data, for the Romanian banking sector, conclusions could also be drawn regarding the methods of analysis of the large volumes of data used and the types of activities specific to the sector, in which big data technologies are used. (www.bnr.ro)

From the point of view of the methods of analyzing large volumes of data, from the data published on the net, the following structure resulted, for the year 2022: (www.bnr.ro)

- 15% use a complete BigData (BD) solution for collecting and processing large volumes of data, including BigData Analytics (BDA);

- 21% use solutions based on DataWarehouse (DW) data warehouses, analyzed with BDA tools;
- 22% use solutions based on DataWarehouse (DW) data warehouses, combined with classic Business Intelligence (BI) analysis solutions;
- 42% use classic analysis and reporting solutions.

The types of data used in BDA analyzes are structured and mostly come from internal sources. For banks that also collect data from external sources, these come from social networks - Facebook, Instagram, Twitter, Youtube, media sources, etc. (www.bnr.ro)

In terms of the types of algorithms used in analyses, most banks using BDA-type analyzes use only static algorithms (eg Fuzzy Matching, Map Reduce, Clustering, Linear Regression), generally for: aggregating data by customer groups , products, segments, periods; creation of customer groups, based on various criteria (e.g. transaction volume/value, balance level); customer grouping, based on various types of criteria (e.g. product types, scoring, collection stage); pre-termination calculation according to product categories; automatic marking of accounts, depending on the configured business characteristics. (www.bnr.ro)

From the review of publicly available information, it emerged that only a few banks reported the use of dynamic Machine Learning algorithms (Logistic Regression, Random Forests, Decision Tree), generally for (www.bnr.ro):

- automatic transaction screening;
- creating alerts, based on the transactional behavioral profile of customers and compared with risk indicators;
- tracking the performance of your own social media channels vs. of competitors;
- optimizing the performance of the content of the own social media channel;
- operational efficiency;
- the development of statistical models, used in the decision-making process of granting credits.

In BDA analysis models, the existence of the so-called Black Box paradox has been noticed, which involves the use of unidentifiable ML Black Box (BB) algorithms, so that the model can only be visualized in terms of input data and output results/analyses, without to know its internal functions. (www.bnr.ro)

5. A Case Study of Electronic Currency Accounting for a Set of Transactions

Table no. 1 Table of transactions related to the case study

Transactions	Textual description of the transactions
01.01.2024	The Sfx99 company purchases 5,000,000 Certificates of Deposit in Digital Currency (CDDC) on 01.01.2024, at the price of 100 monetary units (m.u.). (Nicolae, 2010).
December 31, 2024	On 31.12.2024 the re-estimation is done and it is found that the price of a CDDC has dropped to 95 m.u.
December 31, 2025	On 31.12.2025, the re-estimation is done again, the CDDC reaching the price to 105 m.u.
December 31, 2027	On 31.10.2027 the company sells the CDDC at the price of 120 m.u.

Note:
Debit = D
Credit = C

Source: Case study data proposed by the author.

Accounting data

(Accounting records made by the author)

Table no. 2 *Calculations at 01.01.2024*

Calculations
Purchase cost: 5,000,000 (Certificates of Deposit in Digital Currency CDDC) x 100 um = 500,000,000 um

Source: Calculations made by the author

Table no. 3 Purchase of Deposit in Digital Currency (CDDC) at 01.01.2024

Account - D	**Account - C**	**Amount**
Certificates of Deposit in Digital Currency – CDDC	*Available electronic money*	500,000,000

Source: Calculations and records made by the author

Table no. 4 Calculations at 31.12.2024

Calculations
Cost of re-estimation = 5,000,000 CDDC x (100 – 95) = 25,000,000 um

Source: Calculations made by the author

Table no. 5 Re-estimation of Deposit in Digital Currency (CDDC) at 31.12.2024

Account - D	**Account - C**	**Amount**
Financial expenses - electronic currency	*Certificates of Deposit in Digital Currency – CDDC*	25,000,000

Source: Calculations and records made by the author

Table no. 6 Calculations at 31.12.2025

Calculations
Revenue from *Certificates of Deposit in Digital Currency – CDDC* re-estimation = = 5,000,000 CDDC x (105 – 95) = 50,000,000 mu

Source: Calculations made by the author

Table no. 7 Re-estimation of Deposit in Digital Currency (CDDC) at 31.12.2025

Account - D	**Account - C**	**Amount**
Certificates of Deposit in Digital Currency – CDDC	*Financial revenues - electronic currency*	50,000,000

Source: Calculations and records made by the author

Table no. 8 *Calculations at 31.12.2027*

Calculations
Downloading the book value of *Certificates of Deposit in Digital Currency – CDDC* = 5,000,000 CDDC x 105 = 525,000,000 mu
Earnings from e-currency = 5,000,000 CDDC x (120 – 105) = 75,000,000 mu
Total amount collected = 5,000,000 CDDC x 120 = 600,000,000 mu

Source: Calculations made by the author

Table no. 9 Sale of Deposit in Digital Currency (CDDC) at 31.12.2027

Account - D	**Account - C**	**Amount**
Available electronic money	*Certificates of Deposit in Digital Currency – CDDC*	525,000,000
Available electronic money	Earnings from e-currency	75.000,000

Source: Calculations and records made by the author

6. Conclusions

Our brief analysis results in a set of ideas regarding the possibilities of digital processing of large volumes of accounting and financial data.

New computer technologies open previously unsuspected perspectives regarding the multi-dimensional processing of input data in the accounting information system of various entities. Localized and real-time computer applications of accounting data are being widely spread, applications that allow overcoming the two-dimensional processing, debit - credit, assets - liabilities specific to accounting in the classic system.

We can appreciate that new research horizons are opening up in the field of accounting, horizons determined by the transition to digital information support, replacing the classic paper supports. The perspectives of capitalizing the accounting information in the financial-accounting reports are open, allowing multi-criteria elaborations of the communication of data from the accounting of entities.

The multi-dimensional reporting of accounting information allows the realization of integrated financial statements, on electronic support, in real time, available in the cloud.

7. References

- International Accounting Standards Board, 2011. *International Financial Reporting Standard 13 Fair Value Measurement*. [online] Available at: http://eifrs.ifrs.org
- International Accounting Standards Board, 2022. *International Financial Reporting Standards (IFRSs)*. [online] Available at: http://eifrs.ifrs.org.
- Nicolae, T., 2010. *Standarde contabile [Accounting standards]*. Constanta: ExPonto Publishing House.
- Nicolae, T. C., 2015. *Contabilitate financiară aprofundată: caiet de studiu individidual: învățământ la distanță(ID) [Advanced financial accounting: individual study book for distance learning (DL)]*. Constanta: Ovidius University Press.
- * * * *Law of the Fiscal Code of Romania*, [online] Available at: www.anaf.ro
- www.arbc.ro
- www.ziare.ro
- www.bnr.ro.
- www.hyperledger.org.
- www.insse.ro