

УДК 657.123

DOI: <https://doi.org/10.32782/2520-2200/2019-6-62>**Ілляшенко К.В.**

кандидат економічних наук, доцент,  
доцент кафедри обліку і оподаткування  
Таврійського державного агротехнологічного університету  
імені Дмитра Моторного

**Illiashenko Kateryna**

Dmytro Motornyi Tavria State Agrotechnological University

## РОЛЬ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ У СУЧАСНОМУ БУХГАЛТЕРСЬКОМУ ОБЛІКУ

### THE ROLE OF DATA MINING IN MODERN ACCOUNTING

У статті розглянуто основні форми інтелектуального аналізу даних та вплив сучасних інформаційних технологій на сучасну вітчизняну економіку як у загальному контексті, так і в розрізі окремих підприємств. Акцентовано увагу на необхідності впровадження інформаційних технологій аналізу. Досліджено різні приклади пошуку інформації у великих масивах даних в умовах глобальної інформатизації. Розглянуто переваги та недоліки методів інтелектуального аналізу Data Mining. Проаналізовано етапи пошуку даних та методи, що використовуються в інтелектуальному аналізі даних. Спроектовано архітектуру інформаційно-аналітичної системи поглибленого аналізу звітності. Доведено, що технологія інтелектуального аналізу даних дає змогу аналізувати великі обсяги складних бухгалтерських даних зі швидкістю, яка необхідна для оперативного реагування на фінансові проблеми.

**Ключові слова:** Data Mining, аналіз, дані, знання, інформація, бухгалтерський облік, звітність.

В статье рассмотрены основные формы интеллектуального анализа данных и влияние современных информационных технологий на современную отечественную экономику как в общем контексте, так и в разрезе отдельных предприятий. Акцентировано внимание на необходимости внедрения информационных технологий анализа. Исследованы различные примеры поиска информации в больших массивах данных в условиях глобальной информатизации. Рассмотрены преимущества и недостатки методов интеллектуального анализа Data Mining. Проанализированы этапы поиска данных и методы, используемые в интеллектуальном анализе данных. Спроектирована архитектура информационно-аналитической системы углубленного анализа отчетности. Доказано, что технология интеллектуального анализа данных позволяет анализировать большие объемы сложных бухгалтерских данных со скоростью, необходимой для оперативного реагирования на финансовые проблемы.

**Ключевые слова:** Data Mining, анализ, данные, знания, информация, бухгалтерский учет, отчетность.

Accounting at the enterprises is accompanied by registration of all details of current activity, documentation of all economic operations. The accumulated data form large unstructured clusters that have no useful application without special processing. Such data complicate the process of finding really necessary information. Therefore, they require special processing; the tools of which should be easy to use, and the results should be specific and understandable. One of the methods of such processing is promising, but new for most domestic enterprises, data mining. This study examines the main forms of data mining and the impact of modern information technologies on the modern domestic economy, both in the General context and in the context of individual enterprises. The article focuses on the need for the introduction of information technology analysis. To analyze the information accumulated in modern databases, the methods must be effective: easy to use, have a significant level of scalability and automatism. This concept is at the heart of modern information technology data mining analysis and knowledge discovery in databases (KDD). Various examples of information search in large data sets in the conditions of global informatization are investigated. The advantages and disadvantages of data mining methods are considered. The main advantage of data mining methods is the ability to discover new knowledge that cannot be obtained by traditional methods of economic analysis. The use of data mining methods is not limited. The stages of data search and methods used in data mining are analyzed: classification, regression, clustering, association, sequential patterns, variance analysis, etc. The introduction of information technologies requires new approaches to the analysis of enterprises, which requires the expansion of their analytical activities and more effective use of modern information methods of information processing. Architecture of information and analytical system of in-depth analysis of reports is engineered. It is proved that Data Mining technology allows analyzing large volumes of complex accounting data with the speed necessary for rapid response to financial problems.

**Key words:** Data Mining, analysis, data, knowledge, information, accounting, reporting.

**Постановка проблеми.** У сучасному інформатизованому суспільстві удосконалення технологій збирання та зберігання даних привело до нагромадження величезних потоків різноманітної інформації. Бухгалтерією підприємств здійснюються реєстрація та документування всіх подробиць поточної діяльності. Накопичені дані утворюють неструктуровані накопичені дані, які без спеціального оброблення не мають корисного використання, навпаки, ускладнюють процес пошуку дійсно потрібної інформації. Отже, такі дані потребують особливого оброблення, інструменти якого мають бути простими у застосуванні, а результати – конкретними та зрозумілими. Одним із методів такого оброблення є перспективний, але новий для більшості вітчизняних підприємств інтелектуальний аналіз даних. Все це обумовлює доцільність і актуальність дослідження.

**Аналіз останніх досліджень і публікацій.** Методологія інтелектуального аналізу даних перебуває на стику економічної та технічної науки. Засновником технології вважається математик і програміст Г. Пятецький-Шапіро [1, с. 49]. Проблематикою інтелектуального аналізу даних, вчасності, вивченням Data Mining займалися багато вчених, серед яких можна виділити А. Барсегяна, В. Дюка, М. Купріянова, Д. Макленнена, А. Самойленка, В. Степаненка, Ч. Танга, І. Холода, Х. Ромерсберга. Серед публікацій, що використовувалися у дослідженні, можна виділити роботи, пов'язані з використанням інтелектуального аналізу даних у вітчизняній економіці, які проводили А. Беккаур, В. Вуйцик, Р. Мачуга, А. Пилипенко.

Проблематика використання методів інтелектуального аналізу даних на вітчизняних підприємствах досі ще мало досліджена, особливо щодо взаємозв'язку з бухгалтерським обліком.

**Метою** статті є дослідження ролі інформаційних методів інтелектуального аналізу даних у розрізі його застосування для потреб бухгалтерського обліку на вітчизняних підприємствах.

**Виклад основного матеріалу дослідження.** Головною складовою прийняття управлінських рішень є інформація. Без неї неможлива діяльність будь-якого підприємства. Нині визнано два способи одержання інформації, а саме документальний та експертний. У першому разі дані отримуються з безлічі інформаційних джерел, таких як документи, бази даних, інформаційні системи. Експертний спосіб припускає використання та структурування знань експерта, фахівця з якоїсь галузі. Також їх називають методами, що спрямовані на використання інтуїції та досвіду фахівців [2, с. 32].

До першої групи в економіці належать поширені методи математичної статистики. Такі методи вирішують великий спектр завдань, однак не дають змоги знаходити та витягати знання з масивів даних. Крім того, високі вимоги до кваліфікації кінцевих користувачів обмежують їх використання.

До другої групи належать так звані експертні системи, що є спеціальними комп'ютерними програмами, що моделюють міркування людини (наприклад, експертна система прийняття рішень в управлінському обліку під час планування витрат, експертна система оцінювання ефективності інвестицій) [3, с. 108].

Інформаційні потоки наскрізь пронизують сучасний світ. Особливостями сучасних вимог до оброблення інформації є величезний обсяг даних та їх різноманіття, що робить неспроможними як статистичні, так й експертні підходи в багатьох практичних аспектах, зокрема економічних, тому для аналізу інформації, накопиченої у сучасних базах даних, методи повинні бути ефективними, тобто простими у використанні, мати значний рівень масштабованості й автоматизму. Така концепція полягає в основі двох сучасних інформаційних технологій аналізу, а саме Data Mining і Knowledge Discovery in Databases (KDD) [4, с. 184].

Технологія Data Mining (інтелектуальний аналіз даних) – це видобуток даних з великих масивів. Цей метод виявляє раніше невідомі дані, нетривіальні, практично корисні й доступні для інтерпретації знань, що необхідні для прийняття рішень у різних сферах людської діяльності, тобто інформація, знайдена в процесі застосування методів Data Mining, повинна бути нетривіальною та раніше невідомою. Знання повинні описувати нові зв'язки між властивостями, прогнозувати значення одних ознак на основі інших [5, с. 161].

Наголосимо на тому, що сфера застосування інтелектуального аналізу даних є необмеженою. Найбільше цей метод підходить комерційним підприємствам, що розпочинають проекти на основі інформаційних сховищ даних (Data Warehousing). Досвід таких підприємств показує, що віддача від використання методу може досягати десятків разів. Підприємства за допомогою подібних інформаційних методів аналізу можуть одержати відчутні переваги перед конкурентами та у стратегічному плануванні тощо [6, с. 100].

Найбільшою перевагою методів інтелектуального аналізу даних є здатність виявлення нових знань, які неможливо одержати звичайними методами економічного аналізу, завдяки розширеним алгоритмам завдань. На наш погляд, саме інтелектуальний аналіз буде найбільш затребуваним на підприємствах найближчим часом.

Добуток знань у базах даних KDD – це послідовність дій, яку необхідно виконати для побудови моделі пошуку знання. Ця послідовність не описує конкретний алгоритм або математичний апарат, не залежить від предметної галузі. Це набір атомарних операцій, комбінування яких дає змогу одержати потрібний розв'язок [7, с. 132].

KDD складається з таких етапів, як підготовка вхідних даних, вибір інформативних ознак, очищення даних, застосування методів інтелектуального аналізу даних, підсумкове оброблення та інтерпретація отриманих результатів.

Процес створення моделі інтелектуального аналізу включає такі етапи:

- засвоєння даних (обстеження первинних даних);
- перевірка даних на повноту;
- аналіз ключових факторів впливу;
- виділення винятків;
- виявлення категорій;
- аналіз сценаріїв;
- прогнозування [8, с. 489].

У центрі цього процесу перебувають методи інтелектуального аналізу даних, які дають змогу виявляти закономірності та знання [9].

1) Класифікація – це віднесення об'єктів (спостережень, подій) до одного із заздалегідь відомих класів.

2) Регресія, зокрема завдання прогнозування, – це установлення залежності безперервних вихідних від вхідних змінних.

3) Кластеризація – це угруповання об'єктів (спостережень, подій) на основі даних (властивостей), що описують сутність цих об'єктів. Об'єкти всередині кластера повинні бути «схожими» один на одного й відрізнятися від об'єктів, що ввійшли в інші кластери. Чим більше схожі об'єкти всередині кластера, тим більше відмінностей між кластерами, тим точніше кластеризація.

4) Асоціація – це виявлення закономірностей між зв'язаними подіями. Прикладом такої закономірності є правило, яке зазначає, що з події X впливає подія Y. Такі правила називаються асоціативними. Вперше це завдання було запропоновано для пошуку типових шаблонів покупок, чинених у супермаркетах, тому іноді його ще називають аналізом ринкового кошика (market basket analysis).

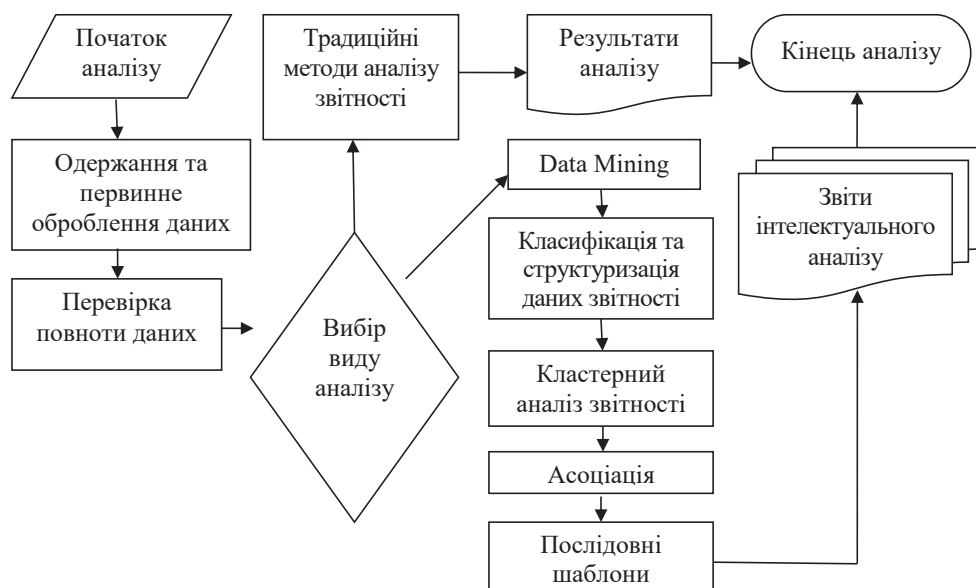
5) Послідовні шаблони – це установлення закономірностей між зв'язаними у часі подіями, тобто виявлення залежності, що якщо відбулась подія X, то через заданий час відбудеться подія Y.

6) Аналіз відхилень – це виявлення найбільш нехарактерних шаблонів.

Так, технологія Data Mining дає змогу аналізувати великі обсяги складних бухгалтерських даних з необхідною для оперативного реагування на фінансові проблеми швидкістю.

Одним з прикладів використання методів інтелектуального аналізу даних у бухгалтерському обліку є виявлення навмисного викривлення даних показників фінансової звітності. Така методика може бути корисною, наприклад, як доказ у судово-бухгалтерській експертизі.

По-перше, створюється історичний профіль підприємства та його керівного складу, який розраховує обсяг операцій з даними бухгалтерського обліку й звітності, їх частоту та пері-



**Рис. 1. Алгоритм аналізу даних фінансової звітності традиційними методами та методами інтелектуального аналізу даних**

Джерело: авторська розробка

одичність. Згодом ці дані структуруються задля зручності аналізу.

По-друге, відбувається кластеризація даних, яку можна проводити у період, відповідний до терміну подання звітності (місячна, квартальна), що є достатнім часовим проміжком для збирання відомостей для аналізу. За вибраними критеріями встановлюються випадки шахрайства, які виходять за межі погрешності більше, ніж на стандартне відхилення, а також відбуваються на нетипові суми з більшою або меншою частотою.

На підставі вищезазначеного нами побудовано алгоритм аналізу даних фінансової звітності, що поєднує традиційні форми економічного аналізу та сучасні методи інтелектуального аналізу даних (рис. 1).

Це не єдиний приклад використання інтелектуального аналізу даних в обліку. Наприклад, цей метод використовується для потреб структуризації інформації відповідно до наявних видів обліку, що є можливим шляхом визначення та правильного встановлення суб'єкта, предмета й об'єктів організації обліково-аналітичного забезпечення [10, с. 199].

Також інтелектуальний аналіз даних можна застосувати в системах стратегічного планування, адже сучасні технології аналізу стають все більш зручними та точними у розрахунках і прогнозуванні, що посилює функції контролю [11, с. 62].

Отже, ми вважаємо, що вже найближчим часом інформаційні системи та інтелектуальні методи аналізу даних будуть використовуватися на більшості вітчизняних підприємств, зокрема під час ведення бухгалтерії.

**Висновки з проведеного дослідження.** Наявний підхід до бухгалтерської інформації має такий недолік, як низький рівень аналітичного оброблення великого обсягу накопичених вторинних даних, тому був запропонований новий підхід з використанням методів оперативного та інтелектуального аналізу даних. Застосування такого методу, як інтелектуальний аналіз, у бухгалтерському обліку може дати змогу підприємствам вийти на новий рівень роботи з інформацією, створити нову парадигму діяльності відповідно до сучасних вимог. Водночас наукова складова цього питання потребує подальших глибоких досліджень.

#### Список використаних джерел:

1. Левкович-Маслюк Л. Великие раскопки и великие вызовы. *Компьютерра*. 2007. № 11 (679). С. 48–51.
2. Knowledge Discovery Through Data Mining: What Is Knowledge Discovery? Tandem Computers Inc. 1996. 54 p.
3. Костякова А. Управлінський облік в інформаційній системі управління витратами. *Економіка АПК*. 2004. № 7. С. 107–110.

4. Hipp J., Guntzer U., Nakaeizadeh G. Algorithms for Association Rule Mining – A General Survey and Comparison. In Proc. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2000. 184 p.
5. Беккауер А. Використання технологій Data Mining для автоматизації бізнес-процесів на виробництві. *Системи обробки інформації*. 2016. № 1 (138). С. 161–164.
6. Ілляшенко К. Сучасні інформаційні методи аналізу даних. *Економічні науки. Серія «Облік і фінанси»*. 2015. № 12 (45). Ч. 2. С. 97–102.
7. Мачуга Р. Методи інтелектуального аналізу наслідків прийняття управлінських рішень як основа перебудови алгоритмів обробки облікової інформації. *Галицький економічний вісник*. 2009. № 2. С. 132–139.
8. Вуйцик В., Смаилова С., Увалиева И. Применение технологии «Data Mining» для анализа данных учебного процесса. *Актуальні проблеми економіки*. 2014. № 1 (151). С. 487–495.
9. Data Mining – добыча данных. URL: <https://basegroup.ru/community/articles/data-mining> (дата звернення: 05.12.2019).
10. Пилипенко А., Пилипенко Д. Розвиток облікової парадигми в контексті сучасних теорій менеджменту та процесів поширення інноваційних знань. *Вісник ЖДТУ*. 2010. № 3 (53). С. 196–200.
11. Голуб Н. Облік і контроль у системі стратегічного планування. *Збірник наукових праць Таврійського державного агротехнологічного університету (економічні науки)*. 2013. № 2 (22). С. 59–65.

#### References:

1. Levkovich-Maslyuk L. (2007) Velikie raskopki i velikie vyzovy. *Kompyuterra*, no. 11 (679), pp. 48–51.
2. Knowledge Discovery Through Data Mining: What Is Knowledge Discovery? (1996) Tandem Computers Inc. 54 p.
3. Kostyakova A. (2004) Upravlinskyi oblik v informatsiinii systemi upravlinnia vytratamy. *Ekonomika APK*, no. 7, pp. 107–110.
4. Hipp J., Guntzer U., Nakaeizadeh G. (2000) Algorithms for Association Rule Mining – A General Survey and Comparison. In Proc. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 184 p.
5. Bekkauer A. (2016) Vykorystannia tekhnolohii Data Mining dlia avtomatyzatsii biznes-protseviv na vyrobnytstvi. *Systemy obrobky informatsii*, no. 1 (138), pp. 161–164.
6. Illiashenko K. (2015) Suchasni informatsiini metody analizu danykh. *Ekonomichni nauky. Seriia "Oblik i finansy"*. 2015. Vol. 2, no. 12 (45), pp. 97–102.
7. Machuha R. (2009) Metody intelektualnoho analizu naslidkiv pryiniattia upravlinskykh rishen yak osnova perebudovy alhorytmiv obrobky oblikovoi informatsii. *Halytskyi ekonomichnyi visnyk*, no. 2, pp. 132–139.
8. Vuysyk V., Smailova S., Uvalieva I. (2014) Primenenie tehnologii "Data Mining" dlya analiza danykh uchebnogo protsessa. *Aktualni problemy ekonomiky*, no. 1 (151), pp. 487–495.
9. Data Mining – добыча данных. Available at: <https://basegroup.ru/community/articles/data-mining> (accessed: 5 December 2019).
10. Pylypenko A., Pylypenko D. (2010) Rozvytok oblikovoi paradyhmy v konteksti suchasnykh teorii menedzhmentu ta protsesiv poshyrennia innovatsiinykh znan. *Visnyk ZhDTU*, no. 3 (53), pp. 196–200.
11. Holub N. (2013) Oblik i kontrol u systemi stratehichnoho planuvannia. *Zbirnyk naukovykh prats Tavriiskoho derzhavnoho ahrotekhnolohichnoho universytetu (ekonomichni nauky)*, no. 2 (22), pp. 59–65.