

УДК 657.123

DOI: <https://doi.org/10.32840/2522-4263/2019-6-68>**Ілляшенко К.В.***кандидат економічних наук,
доцент кафедри обліку і оподаткування
Таврійського державного агротехнологічного університету
імені Дмитра Моторного***Illiashenko Kateryna***Candidate of Economic Sciences,
Associate professor of Accounting and Taxation Department,
Dmytro Motornyi Tavria state agrotechnological University***ВИКОРИСТАННЯ МЕТОДІВ DATA MINING У БУХГАЛТЕРСЬКОМУ ОБЛІКУ****THE APPLICATION OF DATA MINING METHODS IN ACCOUNTING****АНОТАЦІЯ**

У статті розглянуті основні форми інтелектуального аналізу даних та вплив сучасних інформаційних технологій на вітчизняну економіку як у загальному контексті, так і у розрізі окремих підприємств. Акцентовано на необхідності впровадження інформаційних технологій аналізу. Досліджено різні приклади пошуку інформації у великих масивах даних в умовах глобальної інформатизації. Розглянуті переваги та недоліки методів інтелектуального аналізу Data Mining. Проаналізовані етапи пошуку даних та методи, що використовуються в інтелектуальному аналізі даних. Наведено приклад застосування Data Mining для аналізу навмисних помилок у фінансовій звітності. Доведено, що технологія Data Mining дає змогу аналізувати великі обсяги складних бухгалтерських даних зі швидкістю, яка необхідна для оперативного реагування на фінансові проблеми.

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

АННОТАЦИЯ

В статье рассмотрены основные формы интеллектуального анализа данных и влияние современных информационных технологий на отечественную экономику как в общем контексте, так и в разрезе отдельных предприятий. Акцентировано 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. Приведен пример применения Data Mining для анализа преднамеренных ошибок в финансовой отчетности. Доказано, что технология Data Mining позволяет анализировать большие объемы сложных бухгалтерских данных со скоростью, необходимой для оперативного реагирования на финансовые проблемы.

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

ANNOTATION

Work of 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. The article discusses the main forms of data mining and the impact of modern information technologies on the domestic economy, both in the general context and in the context of individual enterprises. The study focuses on the need for the introduction of in-

formation 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. An example of using data mining to analyze intentional errors in financial statements is given. 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, information, accounting, reporting.

Постановка проблеми. Удосконалення технологій збору та зберігання даних у сучасному інформатизованому світі привели до того, що накопичуються величезні потоки різноманітної інформації. Робота підприємств супроводжується реєстрацією та документуванням усіх подробиць поточної діяльності. Накопичені дані утворюють неструктуровані нагромадження, які без спеціальної обробки не мають корисного застосування, більше того, ускладнюють процес пошуку дійсно потрібної інформації. Отже, такі дані потребують особливої обробки, інструменти якої мають бути прості у використанні, а результати повинні бути конкретні та зрозумілі. Одним із методів такої обробки є перспективний, але новий для більшості вітчизняних підприємств інтелектуальний аналіз даних. Все це зумовлює доцільність і актуальність цього дослідження.

Аналіз останніх досліджень і публікацій. Методологія Data Mining знаходиться на стику економічної та технічної науки. Засновником технології вважається американський математик, програміст Г. П'ятецький-Шапіро [1, с. 49]. Проблематикою інтелектуального аналізу даних, зо-

крема 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, І. Чубукову тощо. Серед публікацій, що використовувалися у цьому дослідженні, можна виділити роботи, пов'язані з використанням Data Mining у вітчизняній економіці, які проводили А.О. Беккауер, Р.І. Мачуга, А.А. Пилипенко, О.М. Яцько.

Виділення не вирішених раніше частин загальної проблеми. Проблематика використання методів Data Mining на вітчизняних підприємствах досі ще мало досліджена, особливо у частині взаємозв'язку з бухгалтерським обліком.

Формулювання цілей статті (постановка завдання). Метою статті є дослідження інформаційних методів інтелектуального аналізу даних, вчасності Data Mining, у розрізі його застосування для потреб бухгалтерського обліку на вітчизняних підприємствах.

Виклад основного матеріалу дослідження. Інформація – це головний складник прийняття управлінських рішень, без неї неможлива діяльність сучасного підприємства. Є два способи одержання інформації: документальний і експертний. У першому разі дані отримуються з безлічі інформаційних джерел, таких як документи, бази даних, інформаційні системи тощо. Експертний спосіб припускає використання та структурування знань експерта, фахівця з якоїсь галузі. Часто їх називають методами, що спрямовані на використання інтуїції та досвіду фахівців [2, с. 32].

Серед методів першої групи в економіці поширені методи математичної статистики. Ці 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.

Серед другої групи поширені так звані експертні системи, що являють собою спеціальні комп'ютерні програми, які моделюють міркування людини. Наприклад, експертна система прийняття рішень у управлінському обліку під час планування витрат, експертна система оцінки ефективності інвестицій тощо [3, с. 194]. Висока вартість створення та впровадження експертних систем, нездатність людей виявляти складні нетривіальні залежності, часта відсутність фахівців, що здатні кваліфіковано структурувати свої знання, також заважають поширенню такого підходу [4, с. 73].

Сучасний світ наскрізь пронизаний інформаційними потоками. Особливістю сучасних вимог до обробки інформації є величезний обсяг даних і їх розмаїття, що робить неспроможними як статистичні, так і експертні підходи в багатьох практичних аспектах, у тому числі й економічних. Тому для аналізу інформації, накопиченої у сучасних базах даних, методи повинні бути ефективними: простими у використанні, мати значний рівень масштабованості й автоматизму. Ця концепція полягає в основі двох сучасних інформаційних технологій

аналізу: Data Mining і Knowledge Discovery in Databases (KDD) [5, с. 184].

Технологія Data Mining – «видобуток» даних. Це метод виявлення в «сирих» даних раніше невідомих, нетривіальних, практично корисних і доступних для інтерпретації знань, необхідних для прийняття рішень у різних сферах людської діяльності [6, с. 161]. Тобто інформація, знайдена у процесі застосування методів Data Mining, повинна бути нетривіальною та раніше не відомою. Знання повинні описувати нові зв'язки між властивостями, прогнозувати значення одних ознак на основі інших.

Зауважимо, що сфера застосування Data Mining нічим не обмежена. 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 [7, с. 100].

Основною перевагою методів інтелектуального аналізу Data Mining є здатність виявлення нових знань, які неможливо одержати звичайними методами економічного аналізу. На наш погляд, саме Data Mining буде найбільш затребуваним на підприємствах найближчим часом.

Виявлення 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, комбінуючи які, можна одержати потрібний розв'язок [8, с. 132].

KDD складається з таких етапів: підготовки вхідних даних, вибору інформативних ознак, очищення даних, застосування методів Data Mining, підсумкової обробки даних, інтерпретації отриманих результатів.

Центром усього цього процесу є методи Data Mining, які дають змогу виявляти закономірності та знання [9]:

- класифікація – це віднесення об'єктів (спостережень, подій) до одного із заздалегідь відомих класів;

- регресія, у тому числі завдання прогнозування. Установлення залежності безперервних вихідних від вхідних змінних;

- кластеризація – це угруповання об'єктів (спостережень, подій) на основі даних (властивостей), що описують сутність цих об'єктів. Об'єкти усередині кластера повинні бути «схожими» один на одного і відрізнятися від об'єктів, що ввійшли в інші кластери. Чим більше схожі об'єкти усередині кластера і чим більше відмінностей між кластерами, тем точніша кластеризація;

- асоціація – виявлення закономірностей між зв'язаними подіями. Прикладом такої закономірності є правило, що вказує, що з події X випливає подія Y. Такі правила називаються асоціативними

ми. Уперше це завдання було запропоновано для знаходження типових шаблонів покупок у супермаркетах, тому іноді її ще називають аналізом ринкового кошика (market basket analysis);

– послідовні шаблони – установлення закономірностей між зв'язаними в часі подіями, тобто виявлення залежності, що якщо відбудеться подія X, то через заданий час відбудеться подія Y;

– аналіз відхилень – виявлення найбільш нехарактерних шаблонів.

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

Одним із прикладів використання методів Data Mining у бухгалтерському обліку є виявлення навмисного викривлення даних фінансової звітності. Спочатку створюється історичний профіль підприємства та його управлінського складу, який розраховує обсяг операцій з даними бухгалтерського обліку і звітності, їхню частоту та періодичність. Згодом ці дані структуруються з метою зручності аналізу та застосування.

Наступним етапом є кластеризація даних, яку можна проводити у період, відповідний до терміну подання звітності (місячна, квартальна), це достатній часовий проміжок для збору відомостей для аналізу.

За вибраними критеріями встановлюються випадки шахрайства, які виходять за межі погрішності більш ніж на стандартне відхилення і відбуваються на не типові суми, з більшою або меншою частотою.

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

Сучасні технології аналізу стають все більш зручними та точними у розрахунках і прогнозуванні. Тому ми вважаємо, що вже найближчим часом інформаційні системи та інтелектуальні методи аналізу даних будуть використовуватися на більшості вітчизняних підприємств, у тому числі під час ведення бухгалтерії.

Висновки з цього дослідження. Проблеми сучасної економіки вимагають від вітчизняних підприємств використання лише простих дешевих методів аналізу, яким вважається традиційний економічний аналіз. Але глобальне впровадження інформаційних технологій потребує нових підходів до аналізу діяльності підприємств, яким необхідно розширювати свою аналітичну діяльність та ефективніше застосувати сучасні інформаційні методи обробки інформації. Застосування таких методів, як Data Mining, у бухгалтерському обліку може дати підприємствам змогу створити нову парадигму діяльності відповідно до сучасних вимог. Водночас науковий складник цього питання потребує подальших глибоких досліджень.

БІБЛІОГРАФІЧНИЙ СПИСОК:

1. Левкович-Маслюк Л. Великие раскопки и великие вызовы. *Компьютера*. 2007. 11(679). С. 48–51.
2. Knowledge Discovery Through Data Mining: What Is Knowledge Discovery? Tandem Computers Inc. 1996. 54 p.
3. Безверхня Ю. В. Проблеми та перспективи автоматизації управлінського обліку. *Збірник наукових праць Таврійського державного агротехнологічного університету*. 2015. 1(29). С. 192–198.
4. Яцько О. М. Застосування OLAP-технологій в економічній галузі України. *Науковий вісник Чернівецького університету. Комп'ютерні системи та компоненти*. 2017. Т. 8. Вип. 2. С. 22–30.
5. 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. P. 184.
6. Беккауер А. О. Використання технологій Data Mining для автоматизації бізнес-процесів на виробництві. *Системи обробки інформації*. 2016. 1(138). С. 161–164.
7. Ілляшенко К. В. Сучасні інформаційні методи аналізу даних. *Економічні науки. Серія «Облік і фінанси»*. 2015. 12(45). Ч. 2. С. 97–102.
8. Мачуга Р.І. Методи інтелектуального аналізу наслідків прийняття управлінських рішень як основа перебудови алгоритмів обробки облікової інформації. *Галицький економічний вісник*. 2009. № 2. С. 132–139.
9. Data Mining – добыча данных. URL: <https://basegroup.ru/community/articles/data-mining> (дата звернення: 05.12.2019).
10. Пилипенко А.А., Пилипенко Д.Р. Розвиток облікової парадигми в контексті сучасних теорій менеджменту та процесів поширення інноваційних знань. *Вісник ЖДТУ*. 2010. № 3(53). С. 196–200.

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. Bezverkhnia Yu.V. (2015) Problemy ta perspektivy avtomatyzatsii upravlynskoho obliku. *Zbiryk naukovykh prats Tavriiskoho derzhavnoho ahrotekhnolohichnoho universytetu*, no. 1(29), pp. 192–198.
4. Iatsko O.M. (2017) Zastosuvannia OLAP-tekhnologii v ekonomichnii haluzi Ukrainy. *Naukovyi visnyk Chernivetskoho universytetu. Kompiuterni systemy ta komponenty*, vol. 8, no. 2, pp. 22–30.
5. 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*, p.184 .
6. Bekkauer A.O. (2016) Vykorystannia tekhnolohii Data Mining dlia avtomatyzatsii biznes-protsevis na vyrobnytstvi. *Systemy obrobky informatsii*, no. 1(138), pp. 161–164.
7. Illiashenko K.V. (2015) Suchasni informatsiini metody analizu danykh. *Ekonomichni nauky. Seriya «Oblik i finansy»*. 2015. Vol. 2, no. 12(45), pp. 97–102.
8. Machuha R.I. (2009) Metody intelektualnoho analizu naslidkiv pryiniattia upravlynskykh rishen yak osnova perebudovy alhorytmiv obrobky oblikovoi informatsii. *Halytskyi ekonomichnyi visnyk*, no. 2, pp. 132–139.
9. Data Mining – добыча данных. Available at: <https://basegroup.ru/community/articles/data-mining> (accessed 5 December 2019).
10. Pylypenko A.A., Pylypenko D.R. (2010) Rozvytok oblikovoi paradyhmy v konteksti suchasnykh teorii menedzhmentu ta protsevis poshyrennia innovatsiinykh znan. *Visnyk ZhDTU*, no. 3(53), pp. 196–200.