

Електронний журнал «Ефективна економіка» включено до переліку наукових фахових видань України з питань економіки (Категорія «Б», Наказ Міністерства освіти і науки України № 975 від 11.07.2019). Спеціальності – 051, 071, 072, 073, 075, 076, 292. Ефективна економіка. 2023. № 9.

DOI: <http://doi.org/10.32702/2307-2105.2023.9.21>

УДК 657.1

Б. А. Засадний,

д. е. н., доцент, завідувач кафедри обліку та аудиту,

Київський національний університет імені Тараса Шевченка

ORCID ID: <https://orcid.org/0000-0002-5308-7248>

О. О. Кириллов,

аспірант, Київський національний університет імені Тараса Шевченка

ORCID ID: <https://orcid.org/0009-0000-3770-0156>

ВИКОРИСТАННЯ ІНСТРУМЕНТАРІЮ BUSINESS INTELLIGENCE В АУДИТОРСЬКІЙ ДІЯЛЬНОСТІ

B. Zasadnyi,

*Doctor of Economic Sciences, Associate Professor, Head of Department of
Accounting and Audit, Taras Shevchenko National University of Kyiv*

O. Kyryllov,

Postgraduate student, Taras Shevchenko National University of Kyiv

USING OF BUSINESS INTELLIGENCE TOOLS IN AUDITING

У статті досліджено процес застосування аналітики даних у сучасній аудиторській практиці, що дозволяє аудиторам ефективно використовувати інструменти аналізу, моделювання та візуалізації для виявлення закономірностей, ідентифікації аномалій та отримання корисної

інформації під час аудиту. Обґрунтовано, що сучасні методи *Business Intelligence* та аудиторські інструменти сприяють підвищенню ефективності та достовірності роботи аудиторів. Використання сучасних програмних засобів, таких як *OLAP*-куб та інших інструментів візуалізації, надають можливість аудиторам ефективніше аналізувати великі обсяги даних та отримувати необхідні докази у процесі аудиту. Доведено необхідність впровадження інноваційних методів та інструментів у аудиті для підвищення якості та результативності аудиторської діяльності.

Методи інтелектуального аналізу даних, такі як моделювання, кластеризація та ідентифікація, демонструють високу ефективність у галузі аудиту. В аудиторській діяльності інтелектуальний аналіз даних акцентується на кластеризації, класифікації та прогнозуванні об'єктів перевірки з метою досягнення точної оцінки ймовірності шахрайства. Використання цих методів підвищує точність та швидкість аудиторських процедур, що робить їх більш ефективними і надійними для дослідження діяльності суб'єкта господарювання. У статті досліджено процес використання кластеризації об'єктів для групування даних у відповідності до окреслених аудиторських завдань. Цей підхід допомагає аудиторам зрозуміти закономірності та патерни у великих обсягах даних, що сприяє виявленню потенційних ризиків та можливостей для оптимізації бізнес-процесів клієнта. За результатами кластеризації було виділено 5 груп контрагентів ПрАТ «КиївХліб». Кожен з груп контрагентів було охарактеризовано за допомогою співвідношення між фінансовими показниками. Також було виокремлено потенційних контрагентів для отримання зовнішніх аудиторських доказів з метою перевірки дебіторської заборгованості ПрАТ «КиївХліб».

Доведено, що аналітика даних сприяє розвитку аудиторської професії та допомагає аудиторам більш ефективно виконувати свої обов'язки у сучасному бізнес-середовищі. У статті розглядається актуальність використання сучасних інструментів та технологій аналізу даних, таких як

штучний інтелект і машинне навчання, у процесі аудиторської діяльності. Крім того, у статті надаються практичні поради успішного впровадження аналітики даних в аудиторську практику.

The article researches the process of applying data analytics in modern audit practice, which allows auditors to effectively use analysis, modeling, and visualization tools to identify patterns, identify anomalies, and obtain useful information during an audit. It is substantiated that modern Business Intelligence methods and audit tools contribute to increasing the efficiency and reliability of auditors' work. The use of modern software tools, such as the OLAP cube and other visualization tools, enable auditors to analyze large volumes of data and obtain the necessary evidence in the audit process more efficiently. The need to introduce innovative methods and tools in auditing to improve the quality and effectiveness of audit activity is substantiated.

Methods of data mining, such as modeling, clustering, and identification, demonstrate high efficiency in the field of auditing. In auditing, data mining focuses on clustering, classification, and forecasting of audit objects to achieve precise fraud probability assessments. The utilization of these methods enhances the accuracy and speed of audit procedures, making them more efficient and reliable for examining a business entity's operations.

The article explores the process of using object clustering to group data according to predefined auditing objectives. This approach assists auditors in understanding patterns and trends within large datasets, facilitating the identification of potential risks and opportunities for optimizing client's business processes. As a result of clustering, five groups of counterparties of Kyivkhlіb were identified. Each group of counterparties was characterized based on the relationship between financial indicators. Additionally, potential counterparties for obtaining external audit evidence regarding Kyivkhlіb accounts receivable were singled out.

It is established that data analytics contributes to the advancement of the auditing profession and helps auditors perform their duties more effectively in the contemporary business environment. The article discusses the relevance of using modern data analysis tools and technologies, such as artificial intelligence and machine learning, in the auditing process. Furthermore, practical recommendations for the successful implementation of data analytics in auditing practice are provided.

Ключові слова: *business intelligence, машинне навчання, штучний інтелект, OLAP-куби, кластеризація, метод K-середніх, самоорганізаційна карта Кохонена*

Keywords: *business intelligence, machine learning, artificial intelligence, OLAP cubes, clustering, K-means, self-organizing Kohonen map*

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями. Аналітика великих даних та Business Intelligence є перспективним напрямом для вдосконалення аудиторських перевірок. Застосування методів машинного навчання, аналізу даних та програмного забезпечення візуалізації даних змінюють спосіб роботи аудиторів з їхніми клієнтами. Аудиторські компанії активно використовують програмні інструменти машинного навчання для аналізу договорів, дослідження бухгалтерських кореспонденцій та виявлення фінансових шахрайств. У сфері аналізу даних широко використовуються прогностичні аналітичні інструменти для здійснення прогнозів та оцінок діяльності об'єктів контролю. Інструменти візуалізації даних можуть доповнювати прогностичний аналіз, допомагаючи користувачам виявляти нові тенденції в розвитку бізнес-процесів.

Впровадження системи Business Intelligence в аудиторську діяльність має потенціал кардинально змінити її парадигму та сприяти її удосконаленню. Однією з основних переваг використання Business

Intelligence є можливості обробки великих обсягів даних, а це надає аудиторам можливість аналізувати великі обсяги облікової інформації. Автоматизація процесів збору та аналізу даних завдяки Business Intelligence також зменшує ризик помилок та покращує якість аудиторської роботи. Business Intelligence виступає потужним інструментом для аудиторів, який сприяє підвищенню якості та ефективності аудиторських послуг, а також сприяє їх подальшому розвитку.

Аналіз останніх досліджень і публікацій. Проблематика застосування Business Intelligence у різних галузях економіки розкрита у наукових працях таких вітчизняних вчених, як: М.В. Боровий та Б.В. Дмитришин [1], А.І. Ігнатюк [12], С.В. Івахненко [3], М.В. Шарко та Н.В. Гусаріна [5].

На думку Г.О. Черноус [10], аналітичні програми та бази даних, призначені для впровадження інтелектуального способу управління бізнесом (BI – Business Intelligence), є найбільш перспективними інструментами, запропонованими на ринку технологій обробки та візуалізації даних, які використовуються для прийняття управлінських рішень. Інструменти BI призначені для вимірювання продуктивності бізнесу та оцінки ефективності прийняття управлінських рішень. Виникнення технології BI сприяло створенню систем аналізу інформації нового покоління, які включають різноманітні інструменти видобування даних.

Н. І. Дрокіна, В. Г. Дарчук, О. В. Крижко [2] стверджують, що на сьогодні популяризація та підвищення доступу до BI-інструментів спричинили те, що коло їх користувачів стає значно ширшим. Більшість організацій практикують надання доступу до своєчасної та актуальної інформації окремим категоріям співробітників, відповідальним за прийняття рішень.

О. Г. Яворська [7] зазначає, що бізнес-аналітика сприяє оптимізації даних щодо суб'єктів господарювання та їх капіталізації інтелектуального та інформаційного капіталу, забезпечуючи стійку конкурентну перевагу.

Водночас, необхідно детально розглянути можливості використання інструментарію бізнес-аналітики в рамках проведення аудиторських перевірок, а також визначити їх ступінь ефективності в процесі виявлення фінансових шахрайств.

Формулювання цілей статті. Метою статті є дослідження можливостей використання сучасних методів Business Intelligence у практиці ведення аудиторської діяльності.

Виклад основного матеріалу дослідження. Аудиторські компанії активно тестують можливість використання Business Intelligence у здійсненні аудиторських перевірок. Одним із прикладів виступає використання аудиторською компанією Deloitte програмного рішення Argus. Цей інструмент машинного навчання, який «навчається» з кожної взаємодії з людьми та використовує передові методи машинного навчання та обробки природної мови, щоб автоматично визначати та виділяти ключову облікову інформацію з будь якого типу електронного документа (наприклад, договори оренди, деривативні угоди та договори купівлі-продажу). Програмний продукт Argus дозволяє визначити ключові умови договору, а також виявляти можливі тенденції та відхилення. Це дозволяє не тільки аналізувати договори оренди, визначати ключові терміни, але й інтерпретувати нестандартні договори зі значними судовими вирішеннями (наприклад, ті, які мають незвичайні зобов'язання щодо вибуття активів). Програмний продукт дозволяє аудиторам переглядати та оцінювати більші вибірки (або навіть до 100% усіх первинних документів), витратити більше часу на судові дослідження проблемних аспектів діяльності суб'єкта господарювання та надавати більше інсайдерської інформації, що поліпшує як швидкість, так і якість аудиторського процесу.

Ще одним прикладом технології машинного навчання, яку в даний час використовує компанія PwC є Halo. Інструмент Halo аналізує бухгалтерські проведення та може виявляти потенційно проблемні області (наприклад, записи з ключовими словами сумнівного характеру, записи з

несанкціонованих джерел або надзвичайно велика кількість записів у журналі бухгалтерських проведення, які нижчі за граничні значення). Програмний продукт Halo дозволяє аудиторам перевіряти 100% бухгалтерський проводок та фокусуватися лише на аномальних значеннях з найвищим ризиком викривлення облікової інформації, що суттєво покращує як швидкість, так і якість процедур тестування.

Аналітика даних аудиту (ADA – Audit Data Analytics) допомагає аудиторам виявляти та аналізувати закономірності, ідентифікувати аномалії та вилучати іншу корисну інформацію з даних аудиту за допомогою інструментів аналізу, моделювання та візуалізації. Аудитори можуть використовувати можливості ADA для проведення різноманітних аудиторських процедур зі збору доказів, допомоги при вилученні облікових даних та удосконалення процесу використання аналітики даних, а також для ілюстрації того, як можна використовувати аналітику даних в типовій аудиторській програмі.

Успіх аналітики даних аудиту полягає у використанні сучасного програмного забезпечення, яке спочатку було розроблено для формування баз даних у банківському та роздрібному секторах, а також для забезпечення процесу проектування та моделювання в фінансовому секторі та інжинірингу. Феноменальним у цьому процесі є обсяги даних, які можна ефективно обробляти на промисловому рівні та швидкість обчислень, які виконуються за досить короткий проміжок часу. Тип завдань, які може виконувати таке програмне забезпечення і зв'язки, які можливо встановити, перевищують попередні можливості аудиторського процесу.

За допомогою процедур аналізу даних аудиторські компанії можуть будувати високоякісні статистичні прогнози, які допомагають їм виявляти потенційні ризики, пов'язані з частотою та вартістю господарських операцій. Деякі з цих процедур є простими, інші включають складні моделі. Фахівці, які використовують ці моделі при реалізації аудиторських процедур також

формують професійні судження для визначення математичних та статистичних закономірностей у структурі облікових даних.

Аудитори часто використовують інструментарій аналізу даних для дослідження таких аспектів діяльності суб'єкта господарювання:

- дослідження розрахунків з контрагентами;
- аналіз рентабельності продукції, виокремлення позицій з негативною рентабельністю;
- аналіз капітальних витрат;
- відповідність замовлень та закупівель;
- перевірка бухгалтерських кореспонденцій у журналі господарських операцій.

Незважаючи на те, що методи аналізу даних не можуть повністю замінити традиційні процедури та методи аудиту, але вони виступають потужними інструментами, які дозволяють аудиторам мобільно та якісно здійснювати аудиторські перевірки. Методи аналізу даних надають аудиторам можливості для аналізу усіх можливих транзакцій, які були здійснені суб'єктом господарювання. Таким чином, аудитори можуть потенційно ідентифікувати певний клас сумнівних транзакцій із невідповідними пунктами.

Аудиторські фірми виділяють значні ресурси для дослідження впливу різних методів візуалізації даних на прийняття управлінських рішень. З розвитком технологій та постійним зростанням обсягу даних виникають нові способи представлення облікової інформації. Основною метою візуалізації даних є формування у користувачів уявлення про особливості функціонування суб'єкта господарювання. Це досягається шляхом інтеграції здібностей користувача до процесу аналізу даних та застосування їх гнучкості, креативності до великих наборів даних, які доступні у сучасних системах управління базами даних. Візуалізація даних має кілька основних переваг: лаконічний спосіб представлення інформації та висока швидкість обробки великих даних. Інструменти візуалізації даних є інтуїтивними та не

вимагають розуміння складних математичних або статистичних алгоритмів. Аудитори почали використовувати візуалізацію як інструмент для аналізу кількох рахунків протягом останніх кількох років. Ці інструменти можна використовувати для аналізу аудиторського ризику, тестування транзакцій, аналітичних процедур, підтримки рішень та надання впевненості. Багато рутинних процедур аналізу даних тепер можуть легко виконуватися аудиторами з мінімальним впливом з боку управлінської ланки.

Перспективною функцією інструментів візуалізації даних є їх застосування у комунікації під час аудиту. За допомогою цих інструментів інформацію можна підсумувати та представити способом, який є достатнім для розуміння. Користувач зможе отримати необхідну інформацію за допомогою візуальної презентації. Впровадження технік візуалізації може зробити аналітичні звіти більш зрозумілими, фокусуючись на основних показниках, які аудитор намагається повідомити. Хоча аналіз даних є основою для зовнішнього аудиту, важливо, щоб аудитори розуміли, як працювати з великими обліковими даними. Це допоможе їм краще розуміти особливості функціонування об'єкта контролю та планувати якісний аудит.

Одним з інструментів візуалізації виступає OLAP-куб (Online Analytical Processing Cube) – це багатовимірна структура даних, яка використовується для організації та зберігання інформації з метою її швидкого та зручного аналізу. Він дозволяє користувачам аналізувати дані з різних точок та додаткових вимірювань, надаючи змогу здійснювати зрізи, агрегації та порівняння даних. Аудитори можуть використовувати це для виявлення аномалій, незвичайних транзакцій чи інших надзвичайних ситуацій, які можуть потребувати додаткового обговорення з управлінською ланкою підприємства. Правильне використання OLAP-кубів в аудиторській діяльності може сприяти покращенню якості аудиту, зростанню ефективності та зменшенню ризиків.

За допомогою OLAP-кубу здійснено аналіз процесу реалізації хлібобулочної продукції ПрАТ «КиївХліб». Програмним забезпечення для

побудови OLAP-кубу виступає Loginom Studio. Початковим етапом побудови OLAP-кубу виступає визначення трьох вимірів та ключової міри візуалізатора. Вимірами OLAP-кубу виступає дата реалізації продукції (часовий період – 2022 рік), група товарів (хліб та батони, лаваш, печиво, пряники тощо) та обласний центр постачання (міста України). Мірою для формування OLAP-кубу є вартість реалізації готової продукції. Виходячи із сформованої структури OLAP-кубу було виокремлено 1000 позицій з бази облікових даних ПрАТ «КиївХліб».

Дата	Група товарів	Обласний центр	+	Σ Measures					
		> Конди...	> Круаса...	> Лаваш	> Печив...	> Сухарі ...	> Хліб та ...	Total:	
07.01.2022					7 000	3 000			10 000
08.01.2022			5 600	3 000			7 000		15 600
09.01.2022			3 600	6 600			6 400		16 600
10.01.2022							10 400		10 400
12.01.2022			200	4 000	4 200	4 000	16 400		28 800
13.01.2022		3 000							3 000
14.01.2022				4 400			7 400		11 800
15.01.2022				5 000			12 600		17 600
16.01.2022							12 400		12 400
17.01.2022				4 400			5 200		9 600
18.01.2022			4 800				1 800		6 600
19.01.2022				7 000	3 400	2 000	9 400		21 800

**Рис. 1 OLAP-куб, який характеризує процес реалізації готової продукції
ПрАТ «КиївХліб» за 2022 рік**

[скріншот з програмного продукту Loginom Studio]

Джерело: складено авторами

У програмному продукті Loginom Studio підметом таблиці OLAP-кубу виступає дата реалізації продукції, а присудком таблиці – група товарів та обласний центр постачання. Зміст таблиці характеризує розподіл обсягів реалізації за кожним виміром OLAP-кубу. При цьому по кожному елементу таблиці вказується підсумкова сумарна величина, яка характеризує вартість реалізації продукції за датою постачання, групою товарів та обласним центром.

Одним з ключових аспектів використання OLAP-кубу є формування його зрізів (для цього використовується система фільтрації даних), які

характеризують необхідну для користувача інформацію. Наприклад, необхідна інформація про обсяги реалізації за кожен місяць та квартал (агрегація даних) або оцінка можливостей постачання в рамках одного міста та певної групи товарів.

Дата	Кондитерські вироби	Лаваш	Сухарі та сушки	Хліб та батони	Total:
01.05.2022				3 800	3 800
02.05.2022				7 000	7 000
03.05.2022	600				600
07.05.2022		2 800			2 800
10.05.2022		6 600			6 600
12.05.2022		4 600			4 600
15.05.2022		3 400			3 400
17.05.2022		4 400			4 400
18.05.2022	3 600				3 600
22.05.2022			3 200		3 200
Total:	4 200	21 800	3 200	10 800	40 000

Рис. 2 Зріз OLAP-кубу, який характеризує продаж хлібобулочної продукції ПрАТ «КиївХліб» у місті Києві за травень 2022 року

[скріншот з програмного продукту *Logiном Studio*]

Джерело: складено авторами

Зріз OLAP-кубу дозволяє визначити, що за травень 2022 року у Києві ПрАТ «КиївХліб» реалізувало кондитерських виробів на суму 4200 грн., лавашів – на суму 21800 грн., сухарів та сушок – на суму 3200 грн., а хлібу та батонів – на суму 10800 грн.

OLAP-куби можуть мати різні рівні деталізації та агрегації, що дозволяє користувачам здійснювати аналіз як загальних тенденцій, так і докладних даних. Користувачі можуть обирати конкретні виміри, а також фільтрувати та групувати дані згідно з власними потребами. Використання динамічних методів звітності та візуалізації вирішують проблему великих обсягів облікових даних та дають ефективні результати, які характеризують діяльність суб'єкта господарювання.

Інтелектуальний аналіз даних також виступає одним з найбільш вживаних методів проведення аудиторської перевірки у більшості країн світу. Цей метод використовується для прогнозування ймовірності банкрутства суб'єкта господарювання та виявлення фінансових шахрайств.

Інтелектуальний аналіз даних, який використовується в аудиторській галузі, в основному спрямований на прогнозування. Також наведений інструментарій Business Intelligence використовується в системі управління аудиторськими ризиками, особливо до процесів запобігання ризикам, ідентифікація подій, пом'якшення наслідків негативних подій для підприємства. Коли методи інтелектуального аналізу даних використовуються для виявлення прихованих фактів, які містяться у великих обсягах облікових даних, вони прискорюють процес виявлення фінансових шахрайств. Методи інтелектуального аналізу даних також часто використовуються для аудиту фінансових звітів.

Одним з необхідних елементів для формування аудиторських доказів виступає кластеризація об'єктів залежно від завдань аудиторської перевірки. Кластеризація – це групування об'єктів з урахуванням близькості їх властивостей: кожен кластер складається з подібних об'єктів, а об'єкти різних кластерів значно відрізняються. Кластеризацію використовують, коли відсутні апріорні відомості щодо класів, до яких можна віднести об'єкти досліджуваного набору даних, або коли кількість об'єктів досить велика, що ускладнює ручний процес обробки.

Постановка завдання кластеризації складна і неоднозначна, оскільки:

- оптимальна кількість кластерів у загальному випадку невідома;
- вибір міри подібності чи близькості властивостей об'єктів між собою, як і критерію якості кластеризації, часто має суб'єктивний характер.

Розбиття множини об'єктів групи допомагає виявити внутрішні закономірності, збільшити наочність подання даних, висунути нові гіпотези, оцінити ступінь інформативності властивостей об'єктів.

Для аудиту дебіторської заборгованості ПрАТ «КиївХліб» необхідно здійснити групування цільових контрагентів з метою формування у подальшому аудиторської документації. Виконання цього завдання можливо за допомогою процедур кластеризації, до яких входять ієрархічна кластеризація, метод К-середніх та самоорганізаційні карти Кохонена.

Виконання кластеризації контрагентів ПрАТ «КиївХліб» базується на використанні системи показників, яка включає:

- фактичний термін діяльності (показник надійності контрагентів, який відображає досвід функціонування на ринку);
- обсяг дебіторської заборгованості (ключовий показник аудиторської перевірки);
- коефіцієнт фінансового левериджу (фінансовий коефіцієнт, який відображає співвідношення між власним і позиковим капіталом, тобто оцінює ступінь використання позикових коштів підприємством);
- рентабельність активів (показник, який відображає ефективність використання активів суб'єктом господарювання для генерації прибутку, тобто виступає ідентифікатором ефективності функціонування підприємства).

Інформаційна база даних фінансових показників для кластеризації 30 контрагентів ПрАТ «КиївХліб» представлена у таблиці 1. Для реалізації процедур кластеризації контрагентів ПрАТ «КиївХліб» застосовано програмні продукти STATISTICA та Deductor Studio.

STATISTICA – це програмний продукт, який розроблений компанією StatSoft, а також призначений для аналізу даних, статистичного моделювання, візуалізації і збору інформації різного роду. STATISTICA надає користувачам засоби для виконання широкого спектру аналітичних завдань.

Ієрархічна кластеризація (Tree Clustering) – це сукупність послідовних дій, спрямованих на створення кластерів різного рівня, які підпорядковані між собою відповідно до чітко структурованої ієрархії. Реалізація ієрархічної кластеризації базується на агломеративних (злиттєвих) операціях.

Етапи реалізації ієрархічної кластеризації:

1. Послідовне об'єднання схожих об'єктів з метою створення матриці подібності між об'єктами;
2. Графічна інтерпретація послідовного об'єднання об'єктів у кластери;

3. Створення початкових кластерів із досліджуваної сукупності, а також об'єднання всіх об'єктів в одну велику групу на завершальних етапах аналізу.

Таблиця 1. База даних фінансових показників контрагентів ПрАТ «КиївХліб»

№ контрагента	Фактичний термін діяльності (років)	Обсяг дебіторської заборгованості (тис. грн.)	Коефіцієнт фінансового левериджу	Рентабельність активів
1	20	5890	0,7827	0,0686
2	28	2713	0,2431	0,0820
3	16	5756	0,9436	0,1408
4	8	11833	1,0363	0,1438
5	24	1937	0,3509	0,0957
6	16	6526	0,9112	0,1001
7	26	2767	0,2470	0,0811
8	13	5078	1,0745	0,1415
9	25	1291	0,3359	0,0521
10	23	2402	0,3107	0,0976
11	13	10873	1,9113	0,1281
12	18	6731	0,7931	0,0913
13	22	6594	0,7443	0,0648
14	22	1620	0,7042	0,0843
15	5	5871	3,0490	0,1413
16	24	2744	0,3386	0,0987
17	11	12363	1,9846	0,1193
18	26	1731	0,3428	0,0933
19	13	7075	1,8410	0,1412
20	20	5278	0,9159	0,1204
21	14	6460	0,9935	0,1437
22	29	2577	0,2152	0,0915
23	6	6342	3,2601	0,1386
24	26	2529	0,2028	0,0601
25	10	13109	2,5232	0,1194
26	8	10279	3,3988	0,1011
27	32	2380	0,2290	0,0833
28	12	10520	1,7325	0,1010
29	9	14000	2,1667	0,1227
30	20	5429	0,7874	0,0532

Джерело: складено авторами на основі [13].

Найбільш поширеною мірою оцінки близькості між об'єктами є метрика (або спосіб задання відстані). Проблема вибору тієї чи іншої

метрики є одним з найбільш важливіших у процесі вирішення задач кластеризації.

Найбільш вживаними є такі метрики:

– відстань Евкліда:

$$C_{jk} = \sqrt{\sum_{i=1}^m (X_{ij} - X_{ik})^2} \quad (1)$$

– манхеттенська відстань (відстань Хеммінга):

$$C_{jk} = \sum_{i=1}^m |X_{ij} - X_{ik}| \quad (2)$$

де C_{jk} – відстань між двома об'єктами, X_{ij} – значення ознаки для першого об'єкта, X_{ik} – значення ознаки для другого об'єкта.

Ієрархічна кластеризація може використовуватися для виділення нетипових об'єктів. Це завдання також називають ідентифікацією аномалій (outlier detection). Інтерес ідентифікації представляють кластери (групи), до яких потрапляє мала кількість об'єктів.

Результати ієрархічних процедур кластеризації оформлюються у вигляді деревоподібних діаграм – дендрограм (рис. 3).

Дендрограма допомагає візуалізувати послідовність злиття об'єктів у кластери в процесі ієрархічної кластеризації. Вона надає можливості для розуміння структури ієрархії та відстані між об'єктами на різних рівнях кластеризації.

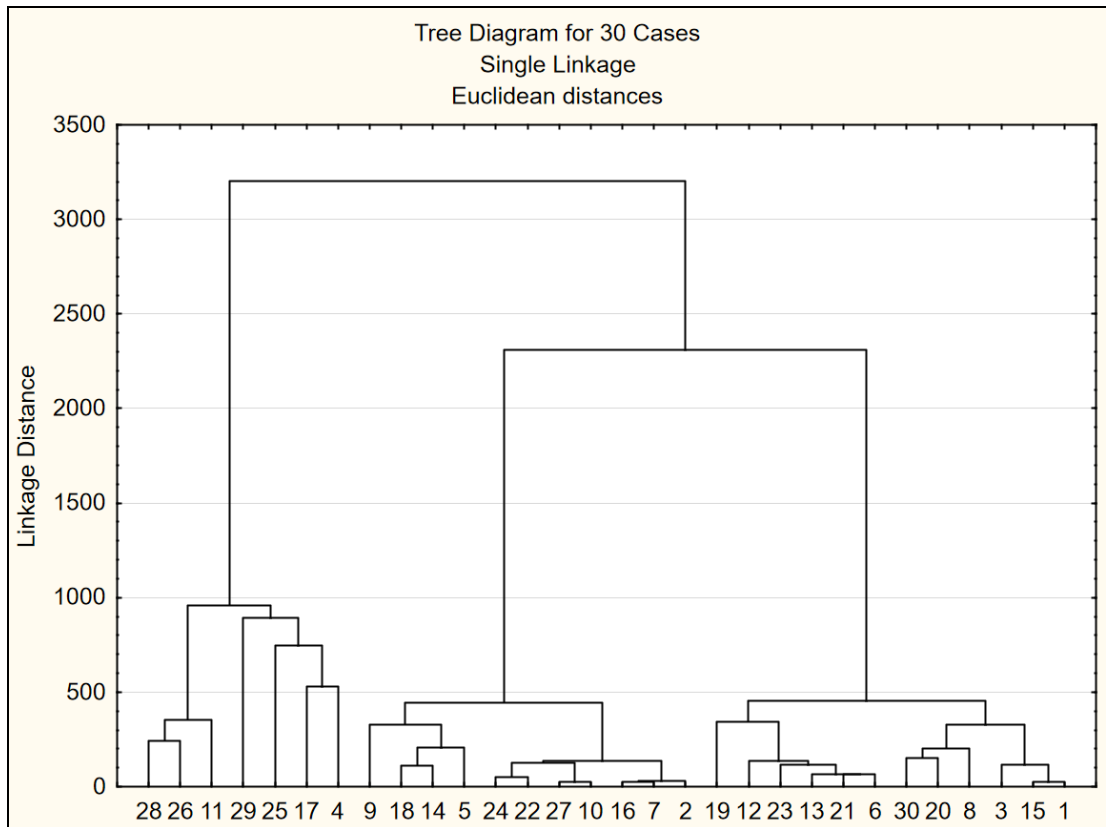


Рис. 3 Дендрограма результатів ієрархічної кластеризації контрагентів ПрАТ «КиївХліб» (оцінка кластерів відбувається за допомогою відстані Евкліда, а формування кластерів здійснюється на основі методу одиничного зв'язку)

[скріншот з програмного продукту STATISTICA]

Джерело: складено авторами

Вертикальна вісь дендрограми зображує відстань об'єднання у кластери, а горизонтальна вісь – контрагентів ПрАТ «КиївХліб». Виходячи зі структури сукупності, до першого кластеру включаються 23,3% усіх контрагентів ПрАТ «КиївХліб», другий та третій кластер містять по 20% від сукупності, четвертий та шостий кластер – по 13,3% сукупності, і тільки п'ятий кластер включає 10% сукупності.

Метод К-середніх (K-means) – це ітераційний алгоритм кластеризації, сутність якого полягає у створенні з первинних даних кластерів одного рангу, ієрархічно не підпорядкованих між собою. Метод К-середніх формує К

кластерів з початкових об'єктів, які розташовані на досить великих відстанях один від одного.

Етапи реалізації методу К-середніх:

1. Вибір кількості кластерів. На початковому етапі кластеризації, обирається кількість кластерів (величина K), на які будуть розподілятися облікові дані.

2. Ініціалізація центрів кластерів. Початкові центри кластерів вибираються випадковим чином або за допомогою встановленого користувачем алгоритму.

3. Призначення центрів до кластерів. Кожний об'єкт призначається до ближчого за відстанню центру кластеру.

4. Розрахунок нових центрів кластерів. Для кожного кластеру обчислюється новий центр як середнє значення всіх об'єктів, які належать до кластеру.

5. Повторна ітерація етапів 3 та 4. Процес призначення точок до кластерів та перерахунку центрів кластерів повторюється декілька разів або до тих пір, поки зміни в кластерах стають незначними або досягається певна кількість ітерацій.

6. Завершення кластеризації. Ітераційні процедури здійснюють до моменту, коли кластерні центри не стабілізуються (на кожній ітерації одні й ті самі об'єкти потрапляють до кожного кластера), дисперсія всередині кластера зменшується до мінімуму, а дисперсія між кластерами – збільшується до максимуму.

Простота та швидкість методу К-середніх обумовлюють відносну популярність використання цього інструментарію для кластеризації різних об'єктів. Проте існують недоліки використання цього методу, до яких можна віднести: чутливість методу до викидів, які наявні у структурі даних; визначення кількості кластерів, залежність результату кластеризації від вибору центрів.

За результатами ієрархічної кластеризації контрагентів ПрАТ «КиївХліб» було визначено, що оптимальним є 6 кластерів для цих об'єктів, тому в процесі реалізації методу К-середніх необхідно сформувати 6 кластерів для контрагентів ПрАТ «КиївХліб». Профілі кластерів контрагентів ПрАТ «КиївХліб» за методом К-середніх представлені на рис. 4.

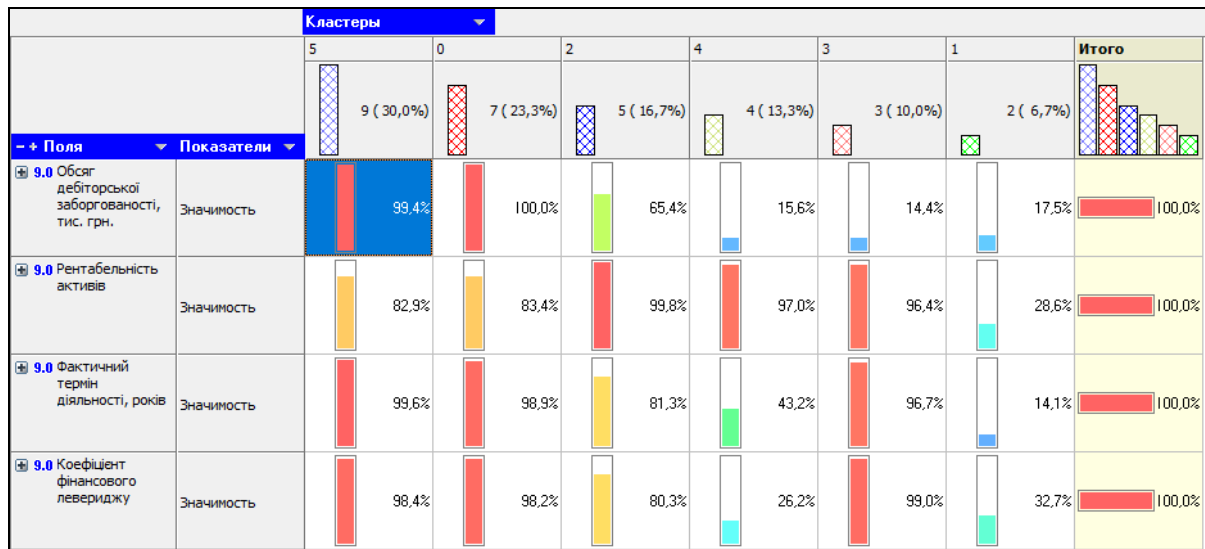


Рис. 4 Профілі кластерів контрагентів ПрАТ «КиївХліб», які отримані за допомогою методу К-середніх

[скріншот з програмного продукту Deductor Studio]

Джерело: складено авторами

За методом К-середніх було сформовано 6 кластерів, перший з яких (кластер 0) містить 7 контрагентів (23,3% сукупності), другий кластер (кластер 1) – 2 контрагентів (6,7% сукупності), третій кластер (кластер 2) – 5 контрагентів (16,7% сукупності), четвертий кластер (кластер 3) – 3 контрагента (10% сукупності), п'ятий кластер (кластер 4) – 4 контрагента (13,3% сукупності) та шостий кластер (кластер 5) – 9 контрагентів (30% сукупності).

Самоорганізаційні карти Кохонена (SOFM – self-organizing feature map) – це нейронна мережа, яка використовується для зведення багатовимірних даних до низьковимірному простору з метою виявлення внутрішньої структури даних і їх кластеризації.

Основна ідея побудови самоорганізаційних карт Кохонена полягає в тому, щоб перетворити вхідні дані у двовимірну сітку нейронів, де кожен нейрон відповідає певному регіону вихідного простору даних. Кожен нейрон має вагові коефіцієнти, які оновлюються під час навчання мережі. Після завершення навчання нейронна мережа може використовуватись для кластеризації нових даних, шляхом знаходження нейрона з найбільшою вагою, який відповідає найбільш схожому класу вхідних даних. Таким чином, самоорганізаційна карта Кохонена може виявляти схожість між даними і групувати їх у певні категорії на основі їх просторового розташування на карті.

Етапи реалізації методу самоорганізаційних карт Кохонена:

1. Ініціалізація вагових векторів. Вагові вектори кожної точки на карті початково встановлюються випадковим чином або використовуються певні підходи для кращої ініціалізації.

2. Вибір вхідного зразка та знаходження центру карти. Обирається один з вхідних зразків даних, на основі яких відбувається пошук точки на карті (центр карти), ваговий вектор якої найбільше схожий на вхідний зразок.

3. Оновлення вагових векторів. Всі вагові вектори навколо центру карти оновлюються таким чином, щоб вони ставали більш схожими на вхідний зразок. Це робиться шляхом зменшення відстані між ваговим вектором і вхідним зразком.

4. Зменшення радіусу вплив. Поступово зменшується радіус, на який впливає центр карти. Це допомагає звужити область оновлення вагових векторів і забезпечити більш глибоке навчання карти.

5. Повторна ітерація етапів 2 – 4. Етапи 2 – 4 повторюються для кожного вхідного зразка декілька разів або до досягнення певної кількості ітерацій.

6. Завершення кластеризації. Після навчання вагові вектори нейронів будуть представляти структуру даних на площині. Кластери можуть бути визначені на основі розташування нейронів на карті.

7. Класифікація нових даних. Після навчання, нові дані можуть бути віднесені до певного кластеру шляхом знаходження нових центрів для кожного нового зразка і включення його до кластера, який найбільше відповідає цьому центру.

Самоорганізаційна карта Кохонена може вимагати значну кількість ітерацій для досягнення збіжності, особливо якщо набір даних досить великий. Під час процесу зведення багатовимірних даних до двовимірної карти може відбутися втрата інформації. Деякі деталі можуть бути втрачені, особливо якщо карта має низьку розмірність.

Самоорганізаційною картою Кохонена було виділено 7 кластерів для контрагентів ПрАТ «КиївХліб» (елемент карти «Кластери»). Профілі кластерів контрагентів ПрАТ «КиївХліб» за методом самоорганізаційних карт Кохонена представлені на рис. 5.

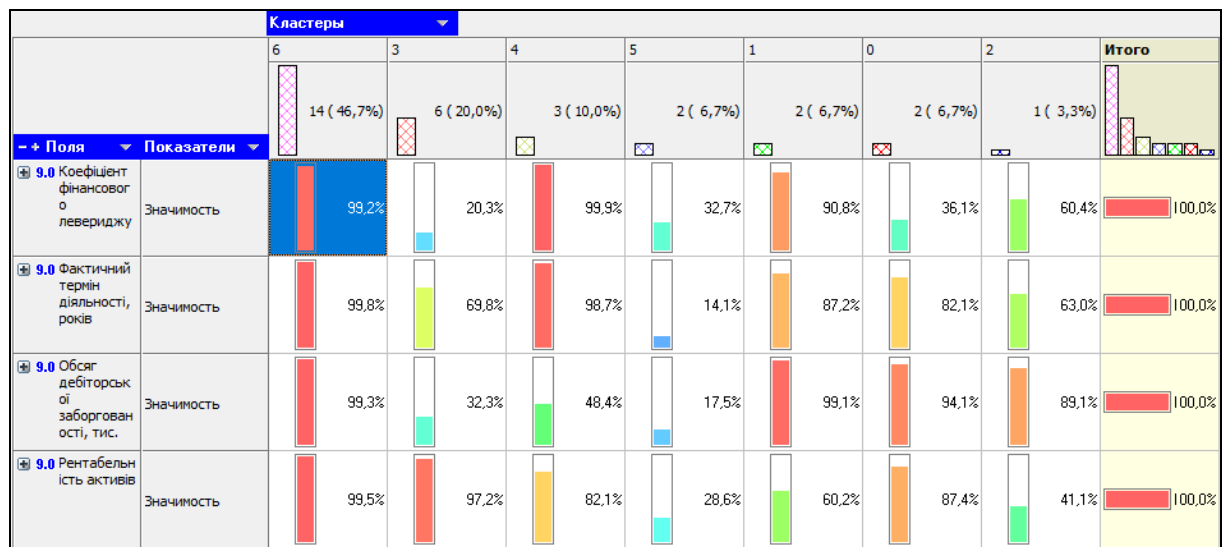


Рис. 5 Профілі кластерів контрагентів ПрАТ «КиївХліб», які отримані за допомогою методу самоорганізаційних карт Кохонена

[скріншот з програмного продукту Deductor Studio]

Джерело: складено авторами

Перший, другий та шостий кластер (кластер 0, 1 та 5) контрагентів ПрАТ «КиївХліб» містить 2 підприємства (6,7% усієї сукупності). Третій кластер (кластер 3) містить найменшу кількість об'єктів – 1 контрагент (3,3%

сукупності). Тому для зручності інтерпретації результатів можливим є об'єднання другого та третього кластеру.

Четвертий кластер (кластер 3) містить 6 контрагентів, що складає 20% від усієї сукупності. П'ятий кластер (кластер 4) включає 3 контрагента (10% сукупності), натомість сьомий кластер (кластер 6) є найбільшим (14 контрагентів – 46,7% від сукупності).

За результатами порівняння було виокремлено 5 груп контрагентів (два кластери було об'єднано в один через відсутність спільних контрагентів за кожним методом). Кожен з груп контрагентів було охарактеризовано за шкалою відібраних показників, крім об'єднання двох кластерів, яке характеризується значним розмахом варіації показників та наявністю викидів у структурі об'єднання кластерів. Тобто контрагенти № 9, 11, 15, 23 та 26 можуть виступати вибірковими об'єктами для отримання зовнішніх аудиторських доказів у цілях подальшої перевірки дебіторської заборгованості ПрАТ «КиївХліб».

Для виявлення шахрайств при веденні бухгалтерського обліку та складанні фінансової звітності обов'язковим стає використання аудиторами інструментарію Business Intelligence, особливо його структурний елемент – інтелектуальний аналіз даних. Інструментарій інтелектуального аналізу даних може бути застосований в межах методології аудиторської діяльності.

Методи інтелектуального аналізу даних продемонстрували високу ефективність у моделюванні, кластеризації та ідентифікації фінансових шахрайств. З точки зору аудиторської діяльності методи інтелектуального аналізу даних зокрема акцентуються на кластеризації, класифікації та прогнозуванні діяльності об'єктів перевірки. Головною метою аудиторської діяльності з використанням інтелектуального аналізу даних є досягнення точної оцінки шахрайства. Методи аудиту, засновані на інструментарії інтелектуального аналізу даних, дозволяють збільшити точність та швидкість аудиторських процедур.

Висновки та перспективи подальших розвідок у даному напрямі.

Аудиторські компанії активно використовують програмні інструменти машинного навчання для аналізу договорів, виявлення фінансових шахрайств та дослідження бухгалтерської кореспонденції. Використання процедур аналізу даних дозволяє аудиторам будувати якісні статистичні прогнози та оцінювати ризики у господарських операціях. Інтеграція інструментів візуалізації даних забезпечує можливість аналізу рахунків, виявлення ризиків та спрощення комунікації під час аудиту. Використання інструментів Business Intelligence та інтелектуального аналізу даних допомагає аудиторам виявляти фінансові шахрайства, покращує точність та швидкість аудиторських процедур.

Для ефективного використання ВІ-інструментів в аудиторській діяльності можуть необхідно розробити спеціальну методику, яка допоможе стандартизувати підходи та покращити якість аудиторських послуг. Подальша розробка ВІ-інструментів, спеціально призначених для задач аудиту сприятиме аудиторам у покращенні ефективності роботи та здійсненню якісного аналізу фінансових даних і виявленню потенційних аудиторських ризиків.

Література

1. Дмитришин Б. В., Боровий М. В. Бізнес-аналітика та її роль в управлінні конкурентоспроможністю підприємства. *Центральноукраїнський науковий вісник. Економічні науки*. 2020. № 5. С. 214-220. URL: http://nbuv.gov.ua/UJRN/Npkntue_2020_5_24.
2. Дрокіна Н. І., Дарчук В. Г., Крижко О. В. Інструменти бізнес-аналітики для візуалізації маркетингових даних. *Причорноморські економічні студії*. 2018. № 26(1). С. 128-138. URL: http://nbuv.gov.ua/UJRN/bses_2018_26%281%29_27.
3. Івахненко С. В. Комп'ютерний аудит. Контрольні методики і технології: навч. посіб. Київ: Знання, 2015. 286 с.

4. Ситник В. Ф., Краснюк М. Т. Інтелектуальний аналіз даних (дейтамайнінг): навч. посіб. Київ: КНЕУ, 2007. 376 с.
5. Шарко М. В., Гусаріна Н. В. Бізнес-аналітика ранжування показників економічної інформації при прийнятті управлінських рішень з інноваційного розвитку виробництва. *Modern economics*. 2018. № 10. С. 146-151. URL: http://nbuv.gov.ua/UJRN/modecon_2018_10_26.
6. Шевченко В. Л., Кірпічніков Ю. А., Федорієнко В. А. Аналіз можливості аналітичної системи SAP Business Intelligence щодо управління оборонними ресурсами. *Збірник наукових праць центру воєнно-стратегічних досліджень Національного університету оборони України*. 2013. № 2. С. 6-13. URL: http://nbuv.gov.ua/UJRN/Znpcvdsd_2013_2_3.
7. Яворська О. Г. Бізнес-аналітика як інструментарій підтримки прийняття рішень в ресторанному бізнесі. *Вісник соціально-економічних досліджень*. 2021. № 3-4 (78-79). С. 174-185. DOI: [https://doi.org/10.33987/vsed.3-4\(78-79\).2021.174-185](https://doi.org/10.33987/vsed.3-4(78-79).2021.174-185).
8. Яремик Х. Я., Яремик М. І. Організація інформаційно-аналітичного забезпечення фінансово-економічної безпеки. *Бізнес Інформ*. 2015. № 10. С. 125-129. URL: http://nbuv.gov.ua/UJRN/binf_2015_10_22.
9. Chu M. K., Yong K. O. Big Data Analytics for Business Intelligence in Accounting and Audit. *Open Journal of Social Sciences*. 2021. Vol. 9. p. 42-52. URL: <https://www.scirp.org/journal/paperinformation.aspx?paperid=111648>.
10. Chornous G., Vetchynov O. Business Intelligence Technology: capabilities, application and features in Ukraine. *Bulletin of Taras Shevchenko National University of Kyiv. Economics*. 2012. Vol. 140. P. 26-29. URL: http://bulletin-econom.univ.kiev.ua/wp-content/uploads/2015/11/140_7.pdf.
11. Chornous G., Farenuk Y., Didenko I. The Guide to Data Mining Course for undergraduate students majoring in Economics. Kyiv: Condor, 2022. 100 p.
12. Ignatiuk A. Business Intelligence for insurance companies. *Bulletin of Taras Shevchenko National University of Kyiv. Economics*. 2016. Vol. 6, No 183. P. 10-15. DOI: <http://dx.doi.org/10.17721/1728-2667.2016/183-6/2>.

13. Сайт Агентства з розвитку інфраструктури фондового ринку України. URL: www.smida.gov.ua.

References

1. Dmytryshyn, B. V. and Borovyj, M. V. (2020), “Business Intelligence and its role in managing a company’s competitiveness”, *Tsentral'noukrains'kyj naukovyj visnyk. Ekonomichni nauky*, [Online], Vol. 5, available at: http://nbuv.gov.ua/UJRN/Npkntu_e_2020_5_24 (Accessed 28 August 2023).
2. Drokina, N. I., Darchuk, V. H. and Kryzhko, O. V. (2018), “Business Intelligence tools for visualizing marketing data”, *Prychornomors'ki ekonomichni studii*, [Online], Vol. 26, No 1, available at: http://nbuv.gov.ua/UJRN/bses_2018_26%281%29_27 (Accessed 28 August 2023)
3. Ivakhnenkov, S. V. (2015), *Komp'uternyj audyt. Kontrol'ni metodyky i tekhnolohii* [Computer audit. Control methods and technologies], Znannia, Kyiv, Ukraine.
4. Sytnyk, V. F. and Krasniuk, M. T. (2007), *Intelektual'nyj analiz danykh* [Data Mining], KNEU, Kyiv, Ukraine.
5. Sharko, M. V. and Husarina, N. V. (2018), “Business Intelligence: ranking economic information indicators in making managerial decisions for innovative production development”, *Modern economics*, [Online], Vol. 10, available at: http://nbuv.gov.ua/UJRN/modecon_2018_10_26 (Accessed 28 August 2023)
6. Shevchenko, V. L., Kirpichnikov, Yu. A. and Fedoriienko, V. A. (2013), “Capabilities of the SAP Business Intelligence analytical system for managing defense resources”, *Zbirnyk naukovykh prats' tsentru voienno-stratehichnykh doslidzhen' Natsional'noho universytetu oborony Ukrainy*, [Online], Vol. 2, available at: http://nbuv.gov.ua/UJRN/Znpcvds_2013_2_3 (Accessed 28 August 2023)
7. Yavors'ka, O. H. (2021), “Business Intelligence as a decision support tool in the restaurant business”, *Visnyk sotsial'no-ekonomichnykh doslidzhen*, [Online],

Vol. 3-4 No 78-79, available at: [https://doi.org/10.33987/vsed.3-4\(78-79\).2021.174-185](https://doi.org/10.33987/vsed.3-4(78-79).2021.174-185) (Accessed 28 August 2023)

8. Yaremyk, Kh. Ya. and Yaremyk, M. I. (2015), “Organization of information and analytical support for financial and economic security”, *Business Inform*, [Online], Vol. 10, available at: http://nbuv.gov.ua/UJRN/binf_2015_10_22 (Accessed 02 September 2023)

9. Chu, M. K. and Yong, K. O. (2021), “Big Data Analytics for Business Intelligence in Accounting and Audit”, *Open Journal of Social Sciences*, [Online], Vol. 9, available at: <https://www.scirp.org/journal/paperinformation.aspx?paperid=111648> (Accessed 3 September 2023).

10. Chornous, G. and Vetchynov, O. (2012), “Business Intelligence Technology: capabilities, application and features in Ukraine”, *Bulletin of Taras Shevchenko National University of Kyiv. Economics*, [Online], Vol. 140, available at: http://bulletin-econom.univ.kiev.ua/wp-content/uploads/2015/11/140_7.pdf (Accessed 28 August 2023).

11. Chornous, G., Farenuk, Y. and Didenko, I. (2022), *The Guide to Data Mining Course for undergraduate students majoring in Economics*, Condor, Kyiv, Ukraine.

12. Ignatiuk, A. (2016), “Business Intelligence for insurance companies”, *Bulletin of Taras Shevchenko National University of Kyiv. Economics*, [Online], Vol. 6, No 183, available at: <http://dx.doi.org/10.17721/1728-2667.2016/183-6/2> (Accessed 28 August 2023).

13. Stock market infrastructure development agency of Ukraine (SMIDA) (2023), [Online], available at: www.smida.gov.ua (Accessed 5 September 2023).

Стаття надійшла до редакції 12.09.2023 р.