

А.І. Пилипенко

КЛАСИФІКАЦІЯ СТАТИСТИЧНИХ ТА МАТЕМАТИЧНИХ МЕТОДІВ АНАЛІЗУ ДАНИХ У ПРОЦЕСІ ПРИЙНЯТТЯ УПРАВЛІНСЬКИХ РІШЕНЬ

Досліджено інтеграцію аналітичних методів аналізу даних в процес прийняття управлінських рішень. Визначено основні характеристики життєвого циклу даних в управлінських рішеннях. Запропоновано класифікацію статистичних та математичних методів за такими ознаками, як час отримання даних (минулі та/або майбутні періоди), досліджувані питання, дії (задачі) на основі аналітичних даних. Рис. 4, табл. 1.

Ключові слова: прийняття управлінських рішень, Big Data, описова аналітика, пізнавальна аналітика, прогностична аналітика, директивна аналітика, попереджувальна аналітика.

JEL C13

Постановка проблеми у загальному вигляді. Кількість даних, які постійно генеруються та розширюються промисловими підприємствами, державними адміністраціями, некомерційними організаціями, науковими установами незмірно збільшується. Ці дані включають текстовий контент (структурований та/або неструктурований), мультимедійний контент (відео, зображення, аудіо) на різноманітних платформах (сайти соціальних мереж, мережі датчиків, кіберсистеми, Інтернет речей та інші). 90% інформації в світі було згенеровано за два останні роки. Зараз всередньому в день генерується близько 2,5 квінтільйонів байт нової інформації [1]. Така кількість складних та неоднорідних даних, що надходять з будь-якого місця, будь-якого часу та будь-якого пристрою, безумовно визначили еру *Big Data*. Сьогодні *Big Data* є стратегічною технологією. Це настільки ж актуально, як нанотехнології та квантові обчислення в сучасній епосі. По суті, *Big Data* є артефактом людського індивідуума, а також колективним інтелектом, який генерується і поширюється головним чином через технологічне середовище, де практично все, що може бути документально зафіксовано, виміряно та сфотографовано в цифровій формі, і таким чином перетворюється на дані [2].

Відповідно до концепції *датафікації* (цей термін з'явився у 2013 році і став все частіше з'являтися у аналітичній літературі [3-5]) та постійно зростаючих технологій, фахівці стверджують, що в майбутньому більша частина даних буде генеруватися та передаватися через машини, оскільки машини спілкуються одна з одною через мережі передачі даних [6]. Незалежно від того, як згенеровані або отримані дані, виникає завдання аналізувати їх таким чином, щоб отримати велику цінність (*Big Value*, цей термін також набуває поширення [7, 8]). Сьогодні *Big Data*, які містять *Big Value*, розглядається як сучасне цифрове паливо (*Digital Oil*), нова сировина XXI-го століття [2]. Відповідна обробка та управління даними може відкрити нові знання та сприяти своєчасному реагуванню на нові можливості та проблеми. Завдяки наявності вдосконалених технологій аналізу *Big Data* (наприклад, NoSQL Databases, BigQuery, MapReduce, Hadoop, WibiData and Skytree) стає можливим краще зрозуміти, як вдосконалити бізнес-стратегії та

процес прийняття рішень у таких важливих сферах, як охорона здоров'я, економіка, енергетика, прогноз природних катаклізмів та ін.

Аналіз останніх досліджень, в яких започатковано розв'язання даної проблеми, і виділення невирішених раніше частин. Опубліковано багато статей, присвячених *Big Data*. Більшість з них є науково-дослідницькими, аналітичними за своїм характером, такими, що зосереджені на використанні експериментів, моделювань, алгоритмів та методів математичного моделювання. Незалежно від їх дослідницького підходу, в цих статтях представлено *Big Data* як джерело нових знань, пропонуючи таким чином інноваційні та дієві підходи для бізнесу [9]. Одночасно з великими можливостями *Big Data* надає великі виклики через безліч джерел з різних доменів. Наприклад, можливості включають багатий бізнес-аналіз для більш інформативних бізнес-рішень [10], підтримка у посиленні видимості та гнучкості ланцюжка постачання та розподілу ресурсів [11] та інше. З іншого боку, такі важливі проблеми, як складність інтеграції даних, брак кваліфікованих кадрів, питання безпеки та конфіденційності даних, синхронізація великих даних ще залишаються викликами для вчених та практиків [12].

У публікації [2] зазначено, що прогрес, досягнутий в сфері *Big Data* та *Big Data Analytics*, не лише виявив відсутність управлінських досліджень у цій галузі, але і відсутність теоретичних конструкцій та академічної суворості, що є функцією методологічної задачі. Загалом, також не вистачає досліджень, які комплексно вирішують основні проблеми *Big Data* або досліджують можливості для нових теорій або нових практик для прийняття управлінських рішень.

Мета статті. Метою статті є класифікація статистичних та математичних методів аналізу даних у процесі прийняття управлінських рішень. Така класифікація спрямована на розширення сфери застосування методів *Big Data*, оцінку різних типів практичних проблем, подолання сучасних викликів в управлінні.

Виклад основного матеріалу дослідження. Концепція *Big Data* є проблематичною для точного визначення, тому що набір даних, який сприймається великим сьогодні, майже неминуче виявиться малим у найближчому майбутньому. Практикуючі фахівці стверджують, що великі набори даних не завжди складні, а малі набори даних не завжди прості, тим самим підкреслюючи, що складність набору даних є важливим фактором у визначенні того чи велика вона. Підтвердженням тому є зміна парадигми від V3 до V5 (рис. 1).

Хоча переваги *Big Data* є фактичними та суттєвими, залишається безліч викликів, які необхідно вирішувати. Деякі з цих завдань пов'язані з функціональними характеристиками *Big Data*, деякі - з існуючими методами та моделями аналізу, а інші - з обмеженням наявної системи обробки. Крім того управлінські рішення, засновані на даних, вимагають аналізу людиною, особою, яка приймає рішення. Незважаючи на те, що обчислювальні технології, необхідні для полегшення цього процесу, йдуть в ногу з часом, людські знання та таланти лідерів бізнесу відстають, і це ще одна велика проблема. Загальні проблеми *Big Data* можуть бути згруповані за трьома основними категоріями на основі життєвого циклу даних (дані, процеси та проблеми управління) [2]:

Задачі пов'язані з характеристиками даних: обсяг даних, різноманітність, швидкість, правдивість, волатильність, якість.

Задачі пов'язані з процесами обробки даних: як зафіксувати дані, як інтегрувати дані, як перетворювати дані, як вибрати правильну модель для аналізу та як надати результати.

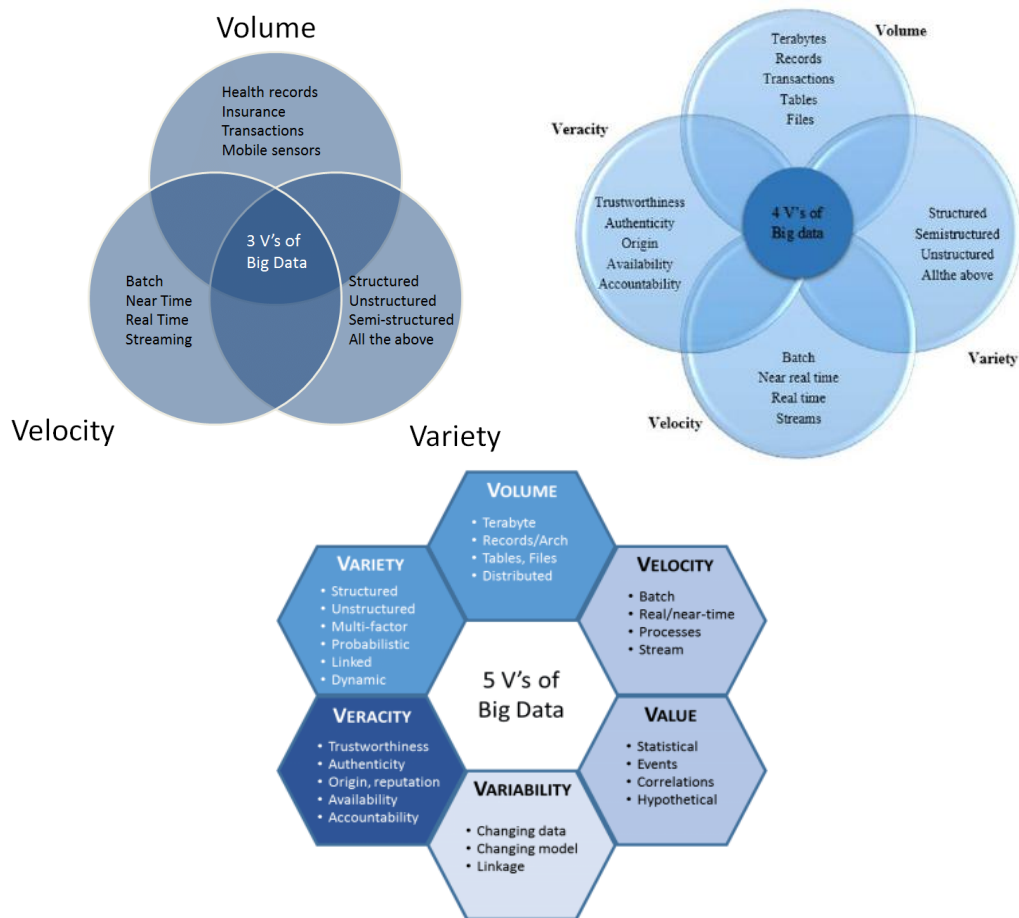


Рис. 1. Еволюція парадигм *Big Data* від V3 до V5 [13-15]

Задачі, пов'язані з управлінням на основі даних: конфіденційність, безпека, управління та етичні аспекти. На рис. 2 представлено часову послідовність зазначених задач в процесі прийняття управлінського рішення на основі аналізу даних.

Щоб полегшити процес прийняття рішень на основі фактичних даних, організаціям потрібні ефективні методи обробки великих об'ємів різноманітної інформації. Потенціал використання *Big Data* нескінченний, але обмежений наявністю технологій, інструментів та навичок у бізнеса. Зрозуміло, що для того, щоб на основі аналізу даних приймати ефективні управлінські рішення, потрібно правильно аналізувати та класифікувати відповідні інструменти та підходи, математичні та статистичні методи.

Для вилучення сенсу з наявних даних використовується різні аналітичні методи, такі як:

описова аналітика (ретельно вивчає дані та інформацію для визначення поточного бізнес-стану таким чином, що стають очевидними тенденції розвитку);
пізнавальна аналітика (допомагає відповісти на запитання: «Чому це відбувається в бізнесі?»);
прогностична аналітика (стосується прогнозування та статистичного моделювання для визначення майбутніх можливостей);
директивна аналітика (полягає в оптимізації та рандомізованому тестуванні, дозволяє оцінити, як бізнес покращує рівень своїх послуг при зменшенні витрат);
попереджувальна аналітика (полягає в формуванні запобіжних заходів на небажані події, наприклад, визначити можливі небезпеки та рекомендувати стратегії пом'якшення наслідків).



Рис. 2. Основні характеристики життєвого циклу даних в управлінських рішеннях

Далі розглянемо окремо кожний тип аналітичних методів.

Описова аналітика є найпростішою формою аналізу даних і включає узагальнення та опис шаблонів знань за допомогою простих статистичних методів, таких як *середня величина, мода, медіана, стандартне відхилення, дисперсія та вимірювання частоти* конкретних подій у потоках даних. Часто великі обсяги історичних даних використовуються в описовій аналітиці для виявлення моделей і створення управлінських звітів, що стосуються моделювання минулої поведінки. Описова аналітика, така як звітність, інформаційні панелі, системи показників та візуалізація даних, вже давно широко використовуються і є основними складовими традиційної бізнес-аналітики. За допомогою описової аналітики можна розкрити не тільки те, що вже відбулося, але і моніторити процес або декілька процесів у різний час. Наприклад, подібне застосування може бути корисним для розуміння умов фінансової стійкості підприємства в певний момент часу або порівняння його з іншими підприємствами однієї галузі, або власними показниками в різні моменти часу. У описовій аналітиці існує потреба в тому, щоб аналітики розвивали вміння читати факти з цифр, зв'язати їх з відповідним процесом прийняття рішень та, нарешті, прийняти рішення на основі аналізу даних. Більша частина бізнес-аналізу

завичай є описовою (дослідницькою), і використання описових статистичних методів дозволяє підприємствам знаходити корисні моделі або невідомі зв'язки (*кореляції*), які можуть бути використані для прийняття бізнес-рішень.

Пізнавальна аналітика полягає у зондуванні даних для сертифікації та/або відхиленні бізнес-пропозицій, для перевірки статистичних гіпотез (аналітичні розрахунки, статистичний аналіз, факторний аналіз). Деякі приклади застосування пізнавальної статистики для KPI показників [16]: Чому довгий час очікування у завантажувальній / розвантажувальній частині? Чому кількість повернень знову зросла? Пошкоджені товари: чи є щось не так з певним матеріалом?

Прогностична аналітика – це клас методів аналізу даних, які сконцентровані на прогнозуванні майбутньої поведінки об'єктів та суб'єктів з метою прийняття оптимальних рішень [17]. Прогностична аналітика використовує статистичні методи, методи інтелектуального аналізу даних, теорії ігор, аналізує поточні та історичні факти для співставлення передбачень про майбутні події. В бізнесі прогностичні моделі використовують закономірності, знайдені в історичних та оперативних даних для ідентифікації ризику та можливостей. Моделі фіксують зв'язки серед багатьох факторів, щоб зробити можливою оцінку ризику та можливостей, пов'язаних з конкретним набором вимог, для прийняття рішень про можливі угоди. Існує декілька підходів до розподілу алгоритмів прогнозного аналізу на групи [18-20]. Окремо виділяють лінійні алгоритми, алгоритми засновані на деревах та методи машинного навчання (наприклад, нейронні мережі) (табл. 1). На рис. 3 наведено приклад зручної схеми представлення алгоритмів машинного навчання.

Директивна аналітика використовується для визначення причинно-наслідкового зв'язку між аналітичними результатами та політикою оптимізації бізнес-процесів. Ця аналітика дозволяє визначити ряд можливих дій і спрямувати на їх реалізацію. Таким чином, директивна аналітика оптимізує моделі бізнес-процесів організації, спираючись на інформацію, надану прогнозними аналітичними моделями. Директивна аналітика сприяє постійній еволюції моделей бізнес-процесів. Загалом, директивні рішення допомагають бізнес-аналітикам у прийнятті рішень, визначаючи дії та оцінюючи їх вплив на бізнес-цілі, вимоги та обмеження. Наприклад, симулятори можуть допомогти дати уявлення про можливі варіанти розвитку бізнесу, які пропонуються для реалізації, щоб закріпити чи покращити свою поточну позицію на ринку.

Попереджувальна аналітика застосовується для того, щоб мати можливість вжити запобіжних заходів. Попереджувальна аналітика намагається кількісно оцінити вплив майбутніх рішень, щоб дати поради щодо можливих результатів, перш ніж рішення дійсно буде зроблено. У найкращому разі, попереджувальна аналітика передбачає не тільки те, що станеться, а й чому це відбудеться. Ця аналітика виходять за межі описової та прогностичної аналітики, рекомендуючи один або декілька можливих кроків. По суті, вона прогнозує кілька рішень і дозволяє бізнесу оцінювати ряд можливих результатів на основі їхніх дій. Попереджувальна аналітика використовує комбінацію алгоритмів машинного навчання та чисельного моделювання. Ці алгоритми використовують дані з багатьох різних джерел, включаючи історичні та операційні дані у режимі реального часу.

Надану вище інформацію про п'ять видів аналітики, що забезпечують процес прийняття управлінських рішень, узагальнено на рис. 4.

Таблиця 1

Порівняльна характеристика прогностичних алгоритмів

Типи алгоритму	Назва	Опис	Переваги	Недоліки
Лінійні	Лінійна регресія	Пошук лінії, яка найкращим чином відповідає точкам даних	Легкі для розуміння	Можливе перенавчання моделі
	Логістична регресія	Адаптація лінійної регресії до проблем класифікації		
Засновані на деревах	Дерево рішень	Побудова графу, який використовує розгалуження, щоб відповідати всім можливим результатам рішення	Легкий для розуміння та реалізації	Не часто використовується для прогнозування, оскільки алгоритм недостатньо потужний для складних даних
	Випадковий ліс	Приймає середнє значення багатьох дерев рішень, кожне з яких зроблено за вибіркою даних. Кожне дерево слабше, ніж повне дерево рішень, але їх об'єднання покращує загальну ефективність	Своєрідна «мудрість натовпу». Алгоритм досягає високоякісних моделей. Моделі швидко тренуються	Повільне прогнозування порівняно з іншими алгоритмами. Нелегко зрозуміти передбачення
	Гradientний бустинг	Використовує ще більш слабкі дерева рішень. Вони все більше фокусуються на «важких» прикладах	Високоєфективні	Невеликі зміни в множині ознак або тренувальній множині можуть привести до радикальних змін у моделі. Не легко зрозуміти передбачення
Нейронні мережі	Нейронні мережі	Наслідують поведінку мозку. Нейромережі є взаємопов'язаними нейронами, які передають повідомлення один одному. Глибоке навчання використовує кілька шарів нейронних мереж, покладених один за одним	Може справлятися з надзвичайно складними завданнями	Дуже повільне тренування через велику кількість. Майже неможливо зрозуміти передбачення

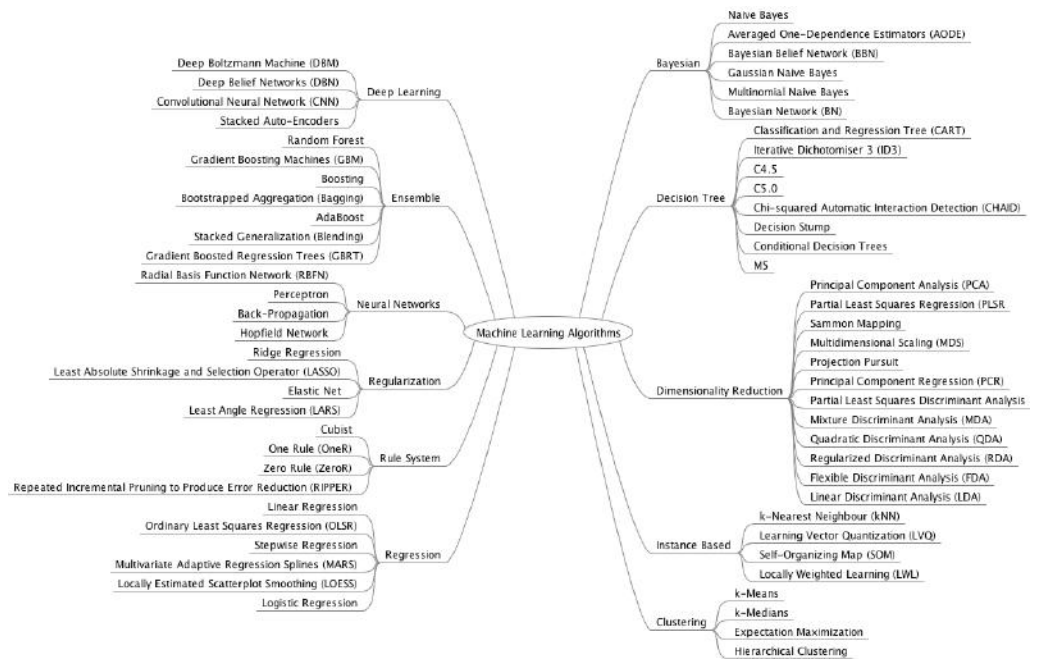


Рис. 3. Приклад класифікації алгоритмів машинного навчання [21]

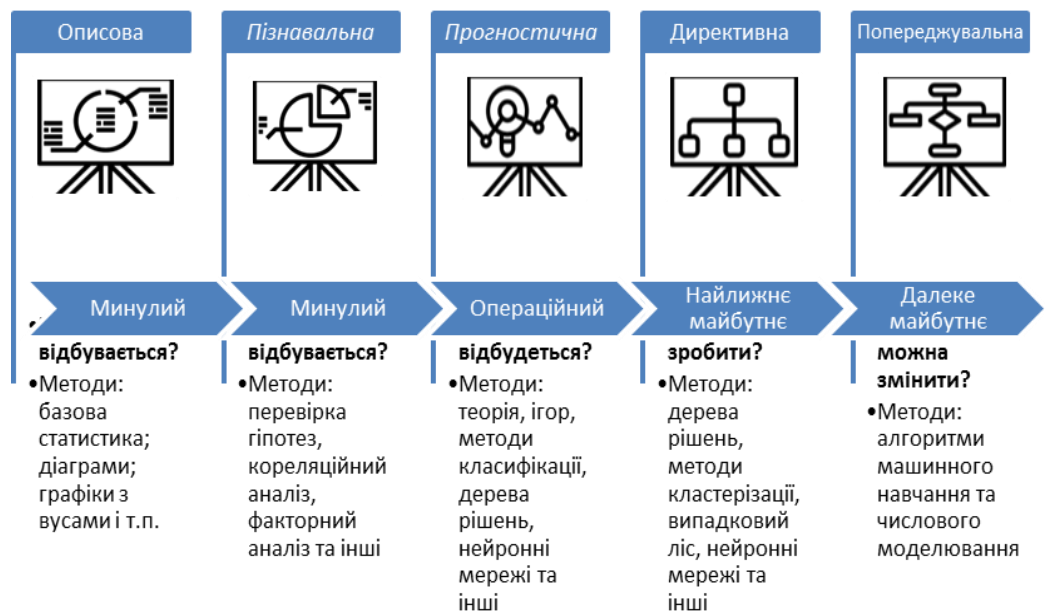


Рис. 4. Класифікація аналітичних методів підтримки прийняття управлінських рішень

Висновки та перспективи подальших досліджень у даному напрямку. Інформація є ключовим фактором успіху, що впливає на ефективність управлінських рішень. Попит на дані постійно зростає, і бізнес розглядає їх як

основний організаційний ресурс для розуміння та прийняття управлінських рішень. Важливим викликом у процесі управління на основі Big Data є неструктурований та складний характер даних. Запропоновано згрупувати загальні проблеми *Big Data* за трьома категоріями: задачі пов'язані з характеристиками даних; задачі пов'язані з процесами обробки даних; задачі пов'язані з управлінням на основі даних. Наведена класифікація статистичних та математичних методів спрямована на те, щоб спростити інтегрування аналітики в процес прийняття рішень, надаючи відповідь на запитання: які дані і які методи застосовувати для вирішення наявних задач.

ЛІТЕРАТУРА

1. How Much Data Does The World Generate Every Minute? // Ifscience. URL: <https://www.ifscience.com/technology/how-much-data-does-the-world-generate-every-minute/>
2. Critical analysis of Big Data challenges and analytical methods. U. Sivarajah et al. // Journal of Business Research 70 (2017) 263–286 URL: <https://www.sciencedirect.com/science/article/pii/S014829631630488X>
3. What Is Datafication and How Can It Impact Your Business? // Import.io URL: <https://www.import.io/post/datafication-impact-business/>
4. Margarita Shilova. The Concept of Datafication; Definition & Examples // Data Science Central URL: <https://www.datasciencecentral.com/profiles/blogs/the-concept-of-datafication-definition-amp-examples>
5. Rahul Zingre. What is Datafication? // LinkedIn URL: <https://www.linkedin.com/pulse/increasing-datafication-our-lives-rahul-zingre>
6. Росс, А. Индустрии будущего / Алек Росс ; [пер. с англ. П. Миронова]. — Москва : Издательство АСТ, 2017. — 351[1] с.
7. Ways to Transform Big Data Analytics into Big Value // Cloudera, Inc. QlikTech International AB. URL: http://cloudera.qlik.com/files/Qlik_Cloudera_BigData_eBook_020818.pdf
8. How Big Data Delivers Big Value to Investors // Institutional Investor LLC. URL: <https://www.institutionalinvestor.com/article/b1968gytts8bm/how-big-data-delivers-big-value-to-investors>
9. Соболева, А.О. Big data: возможности для бизнеса // Научное сообщество студентов: междисциплинарные исследования: сб. ст. по мат. XXXV междунар. студ. науч.-практ. конф. — 2017. — № 24(35). URL: [https://sibac.info/archive/meghdis/24\(35\).pdf](https://sibac.info/archive/meghdis/24(35).pdf)
10. SMART-модель для анализа Big Data в Вашем бизнесе // Business Analysis in Russia. / [Электронный ресурс] – Режим доступа: <https://iiba.ru/smart-model-for-analyzing-big-data-in-your-business/>
11. Видякина, А.О. Большие данные в логистике // Nauka-rastudent.ru. — 2017. — No. 05 (041) URL: <http://nauka-rastudent.ru/41/4258/>
12. Пудов, Д. Проблемы безопасности больших данных // Открытые системы. URL: <https://www.osp.ru/os/2017/04/13053380/>
13. Laney, D. (2001). 3D data management: controlling data volume, velocity, and variety. Meta Group. <http://blogs.gartner.com/douglaney/files/2012/01/ad949-3D-Data-Management-ControllingData-Volume-Velocity-and-Variety.pdf>.
14. Parallel implementation of Apriori algorithms on the Hadoop-MapReduce platform - An evaluation of literature. A.L.Sayeth Saabith, Elankovan Sundararajan, And Azuraliza Abu Bakar // Journal of Theoretical and Applied Information Technology 31st March 2016. Vol.85. No.3 321-351 URL: <https://www.researchgate.net/publication/301544442>
15. Bledi Shaqiri. Exploring Techniques of Improving Security and Privacy in Big Data // Security Techniques in Big Data, 2017. URL: <https://www.researchgate.net/publication/321050765>
16. Descriptive analytics. MAPAD: the new way of thinking // Continuous Improvement Supply Chain Process Management, August 31, 2014. URL: <https://elsekuipers.wordpress.com/tag/descriptive-analytics/>
17. Просчитать будущее: Кто кликнет, купит, совет или умрет / Э. Сигель - М.:Альпина Пабли., 2016. - 374 с.

18. Anubhav Srivastava. Which are the best known machine learning algorithms? Infographic // Think Big Data. URL: <http://thinkbigdata.in/best-known-machine-learning-algorithms-infographic/>
19. Robert Kelley Machine Learning Explained: Algorithms Are Your Friend // Dataiku. URL: <https://blog.dataiku.com/machine-learning-explained-algorithms-are-your-friend>
20. Andreas Mueller. Machine Learning Cheat Sheet (for scikit-learn) // Peekaboo. URL: <http://peekaboo-vision.blogspot.de/2013/01/machine-learning-cheat-sheet-for-scikit.html>
21. Jason Brownlee. A Tour of Machine Learning Algorithms // Machine Learning Mastery URL: <http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>

Рецензент статті
д.т.н., д.е.н., проф. Рамазанов С.К.

Стаття рекомендована до
публікації 19.06.2018 р.

УДК 330.46:65.012.8

О.В. Россошанська

ПРОБЛЕМИ ОЦІНЮВАННЯ ЕКОНОМІЧНОЇ БЕЗПЕКИ СУЧАСНИХ ІННОВАЦІЙНИХ ПРОЕКТНО-ОРІЄНТОВАНИХ ПІДПРИЄМСТВ

За результатами критичного аналізу публікацій за темою дослідження зроблено висновки щодо методологічних підходів та системності в дослідженні економічної безпеки підприємств, а також проблем оцінювання економічної безпеки ІПО підприємств. Показано, що на теперішній час найменше розроблена методологія в економічній безпеці як науковий напряму, не розкрито особливості оцінювання економічної безпеки нового, практично ще не дослідженого економічного об'єкта впливу небезпеки - ІПО підприємств, не враховується нове середовище діяльності для таких підприємств з новими видами загроз – економіка знань. Доведено, що функціонально-діяльнісний системний підхід поки що не був використаний повною мірою в дослідженні систем економічної безпеки, але має великий потенціал для дослідження оцінювання економічної безпеки як наукового напрямку. Зазначені відкриті питання являють собою наукову проблему, яку цілісно представлено у вигляді системної моделі. Рис. 2, табл. 8, дж. 127.

Ключові слова: оцінювання економічної безпеки, підприємство, економіка знань, безпекологія, естиметологія, діяльність, система.

Постановка проблеми у загальному вигляді. У межах кожного наукового напрямку у певні періоди часу виникають методологічні проблеми, які стримують його розвиток. На сучасному етапі однією з таких глобальних проблем у науковому напрямку «Економічна безпека суб'єктів господарської діяльності» є проблема методології оцінювання економічної безпеки взагалі та методологія оцінювання конкретних суб'єктів господарювання. До цього висновку дійшла одна з фундаторів цього наукового напрямку не тільки в Україні, а й у пострадянському просторі, Г. В. Козаченко. В одній з перших праць циклу 2015-2016 років, присвяченому оцінній діяльності в економічній безпекології, вона зазначає наступне: «незважаючи на важливість оцінок економічної безпеки (держави, регіону, підприємства), в економічній безпекології дотепер відсутні цілісні теоретичний та методологічний базиси оцінювання, що призводить до шаблонності оцінювання економічної безпеки підприємства, некоректного використання методу наукової експансії і, як наслідок, низької достовірності оцінок, наявності нечітких правил їхньої інтерпретації і взагалі до деякої імобілізації да стереотипізації оцінювання. ...Взагалі складається враження, що

114 *“Управління проектами та розвиток виробництва”, 2018, №2(66)*