

УДК 004.855

DOI: <https://doi.org/10.53920/ITS-2023-1-1>

**Валерій Вікторович ЗАВГОРОДНІЙ,**

доктор технічних наук, професор,  
завідувач кафедри інформаційних технологій,  
Державного університету інфраструктури та технологій  
ORCID ID: 0000-0002-8347-7183

**Ганна Анатоліївна ЗАВГОРОДНЯ,**

кандидат технічних наук, доцент,  
доцент кафедри інформаційних технологій,  
Державний університет інфраструктури та технологій  
ORCID ID: 0000-0001-8523-1761

**Валентин Олександрович ГОЛОВАЧУК,**

магістр кафедри інформаційних технологій,  
Державний університет інфраструктури та технологій  
ORCID ID: 0009-0004-9001-6512

**ПРОЄКТУВАННЯ СИСТЕМИ РАНЖУВАННЯ  
ПОТЕНЦІЙНИХ КЛІЄНТІВ ТЕЛЕКОМУНІКАЦІЙНОЇ КОМПАНІЇ  
НА ОСНОВІ МАШИННОГО НАВЧАННЯ**

*Робота присвячена проєктуванню системи ранжування потенційних клієнтів телекомунікаційної компанії з огляду на їхню ймовірність придбання продукту на основі машинного навчання, що дозволить оптимізувати бізнес-процес роботи з «холодними» клієнтами.*

*У даній роботі розглядається call-центр телекомунікаційної компанії, спрямований на здійснення вихідних дзвінків з основною метою продажу послуг компанії. Незважаючи на те, що в різних галузях можуть бути власні особливості в діяльності call-центру, загальна модель організації залишається приблизно однаковою, що робить дане дослідження актуальним для будь-якої галузі.*

*Пропонується використовувати історичні дані для навчання алгоритму машинного навчання, який зможе емулювати діяльність супервайзера щодо формування завдань для call-центру.*

*Для кращого розуміння сучасного стану роботи call-центру в роботі наведена діаграма бізнес-процесу формування завдань на дзвінки у нотації BPMN. Також в роботі наведена діаграма, що де-*

**монструє вплив впровадження системи на основі машинного навчання на бізнес-процес формування завдань для обдзвону.**

**Дана робота досліджує задачу ранжування, яка може бути перетворена на задачу бінарної класифікації. В рамках класифікації необхідно визначити ймовірність належності потенційних клієнтів до одного з двох класів, що дозволяє вирішити задачу бінарної класифікації. Перший клас представляє клієнтів, зацікавлених у послугах компанії, тоді як другий клас охоплює клієнтів, які не проявляють інтересу до послуг компанії. У даному контексті найважливішою є ймовірність того, що потенційний клієнт належить до першого класу. Після отримання ймовірностей належності до першого класу, відбувається сортування всіх потенційних клієнтів у порядку спадання ймовірності, вирішуючи таким чином задачу ранжування.**

**Ключові слова:** call-центр, машинне навчання, навчання за прецедентами, система ранжування, аналіз бізнес-процесів, алгоритм класифікації.

**Valerii ZAVGORODNII**

Doctor of technical sciences, Professor,  
Head of the Department of Information Technologies  
State University of Infrastructure and Technologies

**Anna ZAVGORODNYA**

Candidate of technical sciences, Associate Professor,  
Associate Professor of the Department  
of Information Technologies  
State University of Infrastructure and Technologies

**Valentin HOLOVACHUK**

Master of the Department of Information Technologies,  
State University of Infrastructure and Technologies

## **DESIGNING OF THE SYSTEM FOR RANKING POTENTIAL CLIENTS OF A TELECOMMUNICATION COMPANY ON THE BASE OF MACHINE LEARNING**

**The work is devoted to the design of a system for ranking potential customers of a telecommunications company based on their probability of purchasing a product based on machine learning, which will allow optimizing the business process of working with «cold» customers.**

***This work examines the call center of a telecommunications company, aimed at making outgoing calls with the main purpose of selling the company's services. Despite the fact that different industries may have their own peculiarities in the activity of a call center, the general model of the organization remains approximately the same, which makes this study relevant for any industry.***

***It is proposed to use historical data to train a machine learning algorithm that will be able to emulate the activity of a supervisor in the formation of tasks for a call center.***

***For a better understanding of the current state of call center operation, the work presents a diagram of the business process of creating call tasks in BPMN notation. The paper also presents a diagram demonstrating the impact of the implementation of a system based on machine learning on the business process of creating tasks for callback.***

***This work investigates the ranking problem, which can be transformed into a binary classification problem. As part of the classification, it is necessary to determine the probability of potential customers belonging to one of two classes, which allows solving the problem of binary classification. The first class represents customers who are interested in the company's services, while the second class covers customers who do not show interest in the company's services. In this context, the most important thing is the probability that the potential customer belongs to the first class. After obtaining the probabilities of belonging to the first class, all potential customers are sorted in descending order of probability, thus solving the ranking problem.***

***Keywords:*** call center, machine learning, learning by precedents, ranking system, business process analysis, classification algorithm

**Постановка проблеми.** На сьогоднішній день галузь аналізу даних та штучного інтелекту надзвичайно динамічно розвивається, що доводиться щорічним зростанням кількості публікацій з цього напрямку [1]. Помітною є тенденція до зацікавлення штучним інтелектом не лише передових дослідницьких центрів та університетів, але й провідних технологічних компаній, які все частіше створюють власні лабораторії з штучного інтелекту. Цей рівень зацікавленості пояснюється тим, що компанії усвідомлюють великий потенціал методів машинного навчання та активно впроваджують їх у свої бізнес-процеси з метою досягнення значно покращеної ефективності.

Великі компанії проявляють традиційно великий інтерес до методів штучного інтелекту. Вони мають значний обсяг даних для навчання і доступ до ресурсів для дослідження та впровадження таких проєктів. Згідно з аналітичною компанією Gartner, за наступні 2 – 5 років методи машинного навчання та прогнозованої аналітики досягнуть плато продуктивності. Суспільство буде сприймати ці технології та їх впровадження в бізнес-процеси компаній як щось очевидне. Водночас будуть об'єктивно оцінюватися їх можливості, переваги та обмеження [2, 3].

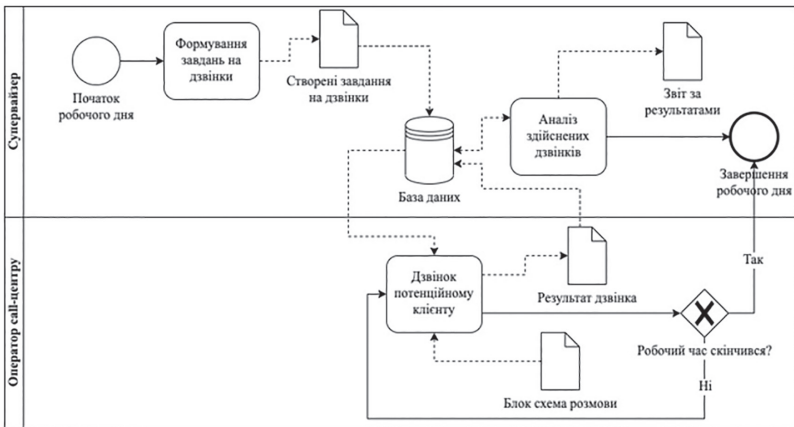
**Аналіз останніх досліджень і публікацій.** Методи машинного навчання вже сьогодні широко використовуються в галузі телекомунікаційних технологій. Телекомунікаційні компанії успішно вирішують багато складних завдань завдяки використанню цих методів. Особливу роль відіграє прогнозна аналітика, яка дозволяє отримати цінну інформацію з великих обсягів даних, які щодня генеруються в цій галузі.

Стаття, опублікована компанією Microsoft [4, 5], засвідчує успішне застосування методів машинного навчання в діяльності call-центрів. В ній приведені реальні приклади завдань, в яких методи машинного навчання здатні повністю або частково замінити людей. Особливістю цієї статті є те, що для кожного описаного завдання наведені ключові показники ефективності (Key Performance Indicators, KPI) та бізнес-цілі, які впливають на ці завдання, а також оцінка ступеня їх впливу. Крім цього, у статті міститься таблиця з описом рекомендованих алгоритмів та обсягів даних, необхідних для вирішення кожного з наведених завдань.

**Мета статті** – проектування системи ранжування потенційних клієнтів телекомунікаційної компанії з огляду на їхню ймовірність придбання продукту на основі машинного навчання, що дозволить оптимізувати бізнес-процес роботи з «холодними» клієнтами.

**Виклад основного матеріалу дослідження.** У даній роботі розглядається call-центр телекомунікаційної компанії, спрямований на здійснення вихідних дзвінків з основною метою продажу послуг компанії. Незважаючи на те, що в різних галузях можуть бути власні особливості в діяльності call-центру, загальна модель організації залишається приблизно однаковою, що робить дане дослідження актуальним для будь-якої галузі.

Одним з важливих елементів будь-якого call-центру, спрямованого на вихідні дзвінки, є супервайзери, чия основна відповідальність полягає у формуванні завдань для операторів. При формуванні цих завдань супервайзери визначають, кому і в якому обсязі слід здійснювати дзвінки від імені call-центру. Це має стратегічне значення, оскільки ефективність роботи всього call-центру залежить від успішного вирішення цього завдання. Для кращого розуміння сучасного стану роботи call-центру була розроблена діаграма бізнес-процесу формування завдань на дзвінки у нотації BPMN [6 – 8]. Зазначена діаграма зображена на рисунку 1.



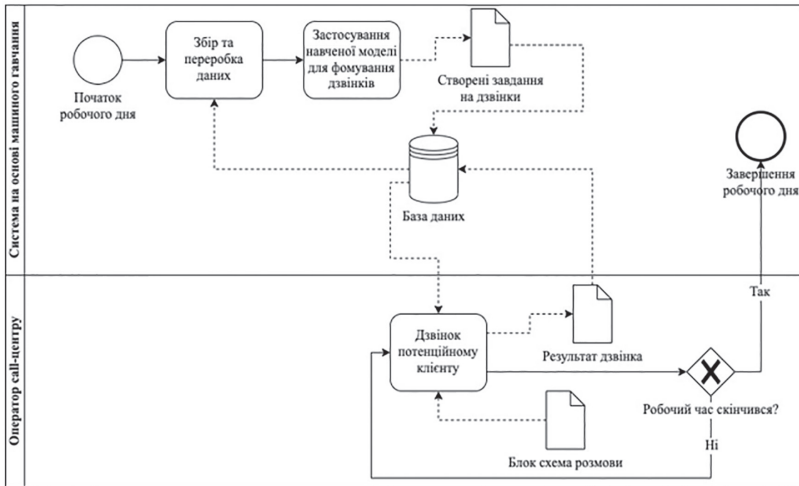
**Рис. 1. Формування завдань на дзвінок у нотації BPMN у стані AS IS**

У бізнес-процесі можна виділити кілька складних етапів. По-перше, щоденно супервайзеру доводиться створювати завдання та аналізувати їх результати після виконання. Засновуючись на аналізі здійснених дзвінків, супервайзер приймає рішення про те, з якими потенційними клієнтами краще зв'язатися з метою збільшення конверсії. Однак виникає проблема, оскільки ретельний та оперативний аналіз результатів є вельми складним завданням через вплив великої кількості факторів. Часто супервайзери мають покладатися на свій досвід та інтуїцію, що не завжди є найоптимальнішим для підвищення ефективності роботи call-центру. Багато часу витрачається на дзвінки до абонентів, які

не зацікавлені у підключенні, що призводить до неефективного використання робочого часу операторів і може шкодити репутації компанії.

Із рисунка 1 видно, що вся інформація про дзвінки, які були створені та оброблені, зберігається у базі даних. Завдяки цьому, використовуючи історичні дані, можна навчити алгоритм машинного навчання, який зможе емулювати діяльність супервайзера щодо формування завдань для call-центру.

Для демонстрації впливу впровадження системи на основі машинного навчання на бізнес-процес формування завдань для обдзвону було підготовлено діаграму у форматі BPMN, яка зображена на рисунку 2.



**Рис. 2. Формування завдань на обдзвін у нотації BPMN у стані TO BE**

Впровадження цього рішення повністю автоматизує процес формування завдань, що виключає необхідність участі супервайзерів. Це означає, що супервайзерам більше не доведеться витрачати час на рутинне формування завдань для call-центру, і вони зможуть приділити більше часу на інші пріоритетні завдання.

Крім того, впровадження системи ранжування може сприяти збільшенню конверсії продажів. Алгоритми машинного навчання

ефективно опрацьовують велику кількість змінних та історичних даних, набагато краще, ніж людський фактор.

Отже, ми маємо задачу спроектувати систему, яка щоденно буде генерувати завдання для телефонних дзвінків у call-центрі з найвищою ймовірністю продажу.

#### *Постановка задачі машинного навчання*

Перед тим, як розпочати проектування системи, важливо чітко визначити, яку задачу машинного навчання ми намагаємося вирішити. Машинне навчання – це галузь штучного інтелекту, що займається вивченням алгоритмів, які можуть самостійно навчатися. Існує кілька типів навчання, але в даній роботі ми будемо використовувати найпоширеніший – навчання за прецедентами.

Цей тип навчання базується на виявленні загальних закономірностей на основі конкретних емпіричних даних. Загалом, задача навчання за прецедентами полягає в тому, що ми маємо навчальний набір даних – скінченну множину прецедентів та їх описи. Наше завдання – виявити залежності, які властиві всім прецедентам, навіть тим, які не входять до навчального набору. Опис кожного прецедента представляється числовим вектором, що складається з ознак, які характеризують всі прецеденти навчального набору. Цей вектор визначає параметри алгоритму. Алгоритм машинного навчання знаходить такі значення параметрів моделі, при яких функціонал якості на навчальному наборі досягає оптимального значення.

Традиційно, існують кілька типів завдань навчання за прецедентами, які перераховані нижче [9, 10]:

1. Завдання класифікації: у цьому випадку необхідно призначити об'єктам одну з обмеженої множини можливих відповідей.

2. Завдання регресії: тут відповіддю є дійсне число, і метою є прогнозування числових значень.

3. Завдання ранжування: це полягає в сортуванні вхідних даних за певним критерієм. Важливо зазначити, що завдання ранжування також може бути сформульоване як завдання класифікації або регресії.

4. Завдання прогнозування: тут об'єктами є відрізки часових рядів, і метою є передбачення майбутніх значень на основі цих рядів.

Дана робота досліджує задачу ранжування, яка може бути перетворена на задачу бінарної класифікації. В рамках класифікації

ми маємо визначити ймовірність належності потенційних клієнтів до одного з двох класів, що дозволить вирішити задачу бінарної класифікації. Перший клас представляє клієнтів, зацікавлених у послугах компанії, тоді як другий клас охоплює клієнтів, які не проявляють інтересу до послуг компанії. У даному контексті нас головним чином цікавить ймовірність того, що потенційний клієнт належить до першого класу. Після отримання ймовірностей належності до першого класу, ми будемо сортувати всіх потенційних клієнтів у порядку спадання ймовірності, вирішуючи таким чином задачу ранжування.

#### *Методи розв'язання задачі класифікації*

Задача класифікації є однією з найбільш вивчених галузей машинного навчання, і вона має велику різноманітність алгоритмів з різними рівнями складності. Ці алгоритми починаються від простих статистичних методів і досягають ресурсоемних методів глибокого навчання.

На сьогоднішній день, серед найпопулярніших і ефективних алгоритмів у всіх галузях класифікації є алгоритми, засновані на деревах та нейронних мережах. Дерева є простими у розумінні інтерпретованими моделями, які добре працюють навіть на невеликих наборах даних. Вони використовують розбиття ознак на дерево рішень, що дозволяє легко визначити класифікаційні правила.

Нейронні мережі, зокрема багат шарові перцептрони, здатні до складних нелінійних моделей та автоматичного вивчення високорівневих ознак. Вони здатні працювати з великими обсягами даних та виявляти складні залежності між вхідними і вихідними даними. Завдяки глибині та ширині нейронних мереж можна досягти вражаючих результатів у багатьох задачах класифікації.

Проте, необхідно враховувати, що існують ситуації, коли прості алгоритми можуть показувати кращі результати в окремих задачах. Це може статися з причини обмеженості даних або недостатньої кількості навчальних прикладів для ефективного навчання складних моделей. Тому варто розглядати широкий спектр алгоритмів і не виключати простих методів, наприклад, таких як наївний Баєс, логістична регресія або метод опорних векторів. Ці алгоритми можуть бути ефективними в ситуаціях, коли дані мають чітку структуру або коли кількість ознак обмежена.

Крім того, прості алгоритми зазвичай мають менші вимоги до обчислювальних ресурсів і можуть працювати швидше ніж склад-



ні моделі. Це особливо важливо, якщо є необхідність обробляти великі обсяги даних в реальному часі або на пристроях з обмеженими обчислювальними можливостями.

Отже, вибір алгоритму для задачі класифікації залежить від конкретного контексту, доступних даних, обчислювальних ресурсів та конкретних потреб. Варто провести експерименти з різними алгоритмами і здійснити порівняльний аналіз результатів, щоб знайти найбільш ефективне рішення для конкретної задачі класифікації.

#### *Вимоги до системи ранжування*

Під час проектування системи ранжування необхідно визначити вимоги, які повинні бути задоволені. Основні вимоги до системи включають наступне:

1. Збір ознакового опису потенційних клієнтів з корпоративної бази даних.

2. Використання навченої моделі машинного навчання для отримання ймовірності продажу.

3. Формування вибірок потенційних абонентів з найвищою ймовірністю продажу (розмір вибірок заздалегідь визначений).

4. Взаємодія з корпоративною системою формування завдань для call-центру.

5. Автоматичне оновлення моделі машинного навчання.

6. Формування звіту про оновлення моделі та його розсилка через корпоративну пошту.

7. Формування звіту про результати роботи моделі та його розсилка через корпоративну пошту.

#### *Загальна архітектура системи ранжування потенційних клієнтів*

На рисунку 3 зображена загальна структура системи у вигляді компонентної діаграми. Розроблювана система взаємодіє з корпоративною базою даних та системою формування завдань для call-центру, що є внутрішньою системою компанії. Остання забезпечує інтерфейс для взаємодії з розроблюваною системою.

Система складається з п'яти компонентів, кожен з яких виконує певні функціональні завдання, що описані нижче:

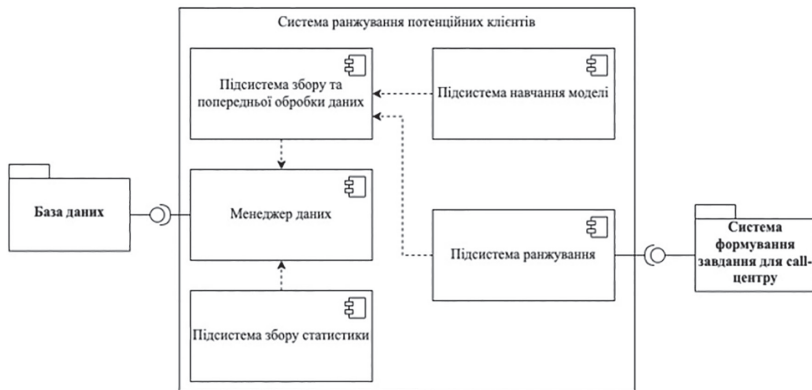
1. Менеджер даних: цей компонент відповідає за взаємодію з корпоративною базою даних і забезпечує основний функціонал обробки даних.

2. Підсистема збору та попередньої обробки даних: ця підсистема відповідає за формування навчальної вибірки та підготовку даних для щоденного ранжування.

3. Підсистема збору статистики: цей компонент виконує збір статистичних показників роботи системи з метою постійного моніторингу якості.

4. Підсистема навчання моделі: ця підсистема відповідає за підтримку навченої моделі машинного навчання в актуальному стані.

5. Підсистема ранжування: цей компонент взаємодіє із зовнішньою системою формування завдань для call-центру і відповідає за щоденне формування вибірок потенційних абонентів з найбільшою ймовірністю продажу.



**Рис. 3. Діаграма компонентів**

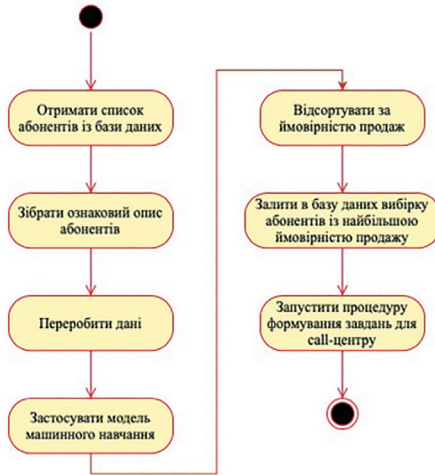
*Основний функціонал системи*

Розроблювана система має виконувати три основні функції, а саме:

1. Щоденне формування завдань для call-центру з максимальною ймовірністю продажу.
2. Оновлення навченої моделі машинного навчання.
3. Розсилка звіту з результатами роботи системи.

Виконання кожної з цих функцій має бути ініційоване операційною системою в наперед встановлений час.

На рисунку 4 представлена загальна схема роботи щоденного процесу формування завдань для call-центру.



**Рис. 4. Загальна схема процесу формування завдань**

З істотним перебігом часу, тенденції на ринку послуг інтернет-провайдерів можуть зазнавати змін. Певні ознаки можуть втратити свою прогностичну вагу, тоді як інші можуть набувати більшої важливості. З цієї причини, необхідно реалізувати функціонал системи, спрямований на підтримку актуальності моделі. На рисунку 5 наведена загальна схема оновлення моделі машинного навчання.



**Рис. 5. Загальна схема оновлення моделі**

Для моніторингу ефективності системи та своєчасного прийняття рішень, необхідно врахувати можливість збору статистики і формування звіту про результати роботи системи. Також, важливим є автоматична відправка цього звіту через корпоративний поштовий сервіс. Схема роботи, зображена у формі діаграми діяльності, представлена на рисунку 6.



**Рис. 6. Схема роботи процесу формування звіту**

Процес надсилання звітів може бути повторюваним, що дозволяє системі регулярно надсилати звіти про роботу моделі машинного навчання. Звіти, що відправляються, можуть бути використані для аналізу результатів, моніторингу продуктивності та прийняття рішень щодо подальшого розвитку моделі.

**Висновки та пропозиції.** Основна ціль роботи полягала в проектуванні системи, яка могла б ранжувати потенційних клієнтів залежно від ймовірності здійснення продажу. Це дозволило б не тільки замінити рутинну роботу супервайзерів під час формування завдань для call-центру, але й підвищити конверсію продажів. Для досягнення цієї цілі були виконані такі завдання:

- проведено аналіз бізнес-процесу формування завдань для call-центру;
- проведено порівняльний аналіз алгоритмів класифікації;
- спроектована система ранжування.

У подальшому розвитку планується виконати наступні кроки:

1. Провести розробку нових ознак, які зможуть якісно описати потенційних клієнтів компанії.

2. Здійснити експерименти з новими алгоритмами класифікації з метою покращення якості метрик.

3. Розробити зручний веб-інтерфейс для управління системою ранжування.

Ці кроки спрямовані на поліпшення функціональності та ефективності системи, забезпечуючи більш точний опис потенційних клієнтів, покращення класифікаційних алгоритмів та зручний інтерфейс для їх керування.

© **Завгородній В.В., Завгородня Г.А., Головачук В.О., 2023**

## ЛІТЕРАТУРА

1. Karen Hao. We analyzed 16,625 papers to figure out where AI is headed next. URL: <https://www.technologyreview.com/2019/01/25/1436/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/> (дата звернення: 30.05.2023).

2. Eka Ponkratova. Finally available and accessible, but wrong way? URL: <https://towardsdatascience.com/finally-available-and-accessible-but-wrong-way-bec731a151a6> (дата звернення: 30.05.2023).

3. Gartner Reveals. Five Major Trends Shaping the Evolution of Analytics and Business Intelligence. URL: <https://www.gartner.com/en/newsroom/press-releases/2019-10-02-gartner-reveals-five-major-trends-shaping-the-evoluti> (дата звернення: 30.05.2023).

4. Business Decision Makers. Machine learning implementation strategy for a customer service center. URL: <https://cloudblogs.microsoft.com/dynamics365/bdm/2018/02/07/machine-learning-implementation-strategy-for-a-customer-service-center/> (дата звернення: 30.05.2023).

5. Gartner Predicts. The Future Of AI Technologies. URL: <https://www.gartner.com/smarterwithgartner/gartner-predicts-the-future-of-ai-technologies> (дата звернення: 30.05.2023).

6. Pankaj E. Kasar. BPMN To REO Modeling Transformation. LAP LAMBERT Academic Publishing. 2017. P. 52. ISBN 9783330072145.

7. Vivek Kale. Enterprise Process Management Systems. CRC Press, 2018. P. 437. ISBN 9780429842337.

8. Philippe Desfray, Gilbert Raymond. Modeling Enterprise Architecture with TOGAF. Morgan Kaufmann. 2014. P. 350. ISBN 9780124199958.

9. Matthew Mayo. Top Data Science and Machine Learning Methods Used in 2018, 2019. URL: <https://www.kdnuggets.com/2019/04/top-data-science-machine-learning-methods-2018-2019.html> (дата звернення: 30.05.2023).

10. Royce Reinger. A Game Theoretic Approach to Explain The Output Of any ML Model. URL: <https://morioh.com/p/70356b3ecbc8> (дата звернення: 30.05.2023).

## REFERENCES

1. Hao, K. (2019), «We analyzed 16,625 papers to figure out where AI is headed next», available at: <https://www.technologyreview.com/2019/01/25/1436/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/> (Accessed 30 May 2023).

2. Ponkratova, E. (2020), «Finally available and accessible, but wrong way?», available at: <https://towardsdatascience.com/finally-available-and-accessible-but-wrong-way-bec731a151a6> (Accessed 30 May 2023).

3. Gartner Reveals (2019), «Five Major Trends Shaping the Evolution of Analytics and Business Intelligence», available at: <https://www.gartner.com/en/newsroom/press-releases/2019-10-02-gartner-reveals-five-major-trends-shaping-the-evoluti> (Accessed 30 May 2023).

4. Business Decision Makers (2018), «Machine learning implementation strategy for a customer service center», available at: <https://cloudblogs.microsoft.com/dynamics365/bdm/2018/02/07/machine-learning-implementation-strategy-for-a-customer-service-center/> (Accessed 30 May 2023).

5. Gartner Predicts (2019), «The Future Of AI Technologies», available at: <https://www.gartner.com/smarterwithgartner/gartner-predicts-the-future-of-ai-technologies> (Accessed 30 May 2023).

6. Kasar, P.E. (2017), «BPMN To REO Modeling Transformation». LAP LAMBERT Academic Publishing. P.52. ISBN 9783330072145.

7. Kale, V. (2018), «Enterprise Process Management Systems». CRC Press. P.437. ISBN 9780429842337.

8. Desfray P., Raymond G. (2014), «Modeling Enterprise Architecture with TOGAF». Morgan Kaufmann. P.350. ISBN 9780124199958.

9. Mayo, M. (2019), «Top Data Science and Machine Learning Methods Used in 2018, 2019», available at: <https://www.kdnuggets.com/2019/04/top-data-science-machine-learning-methods-2018-2019.html> (Accessed 30 May 2023).

10. Reinger, R. (2023), «A Game Theoretic Approach to Explain The Output Of any ML Model», available at: <https://morioh.com/p/70356b3ecbc8> (Accessed 30 May 2023).

**СТАТТЯ НАДІЙШЛА ДО РЕДАКЦІЇ 23.05.2023**