

УДК 342.95

**МОДЕЛІ МАШИННОГО  
НАВЧАННЯ ДЛЯ ІНФОРМАЦІЙНО-  
ПРАВОВОГО ЗАБЕЗПЕЧЕННЯ  
ДІЯЛЬНОСТІ СУДУ****Ольга КОВАЛЬЧУК**

кандидат фізико-математичних наук,  
доцент, доцент кафедри теорії права та  
конституціоналізму

Західноукраїнський національний  
університет (Тернопіль)

<https://orcid.org/0000-0001-6490-9633>

Алгоритми штучного інтелекту на основі систем машинного навчання вже сьогодні використовують для інформаційно-правової підтримки ухвалення ефективних судових рішень у високотехнологічних країнах світу. Україна також розпочала цифрову трансформацію судочинної системи. Формування надійного інформаційного забезпечення судів потребує використання сучасних методів науки про дані та інноваційних інформаційних технологій для якісного аналізу великих обсягів даних.

У статті представлено скорингову модель для оцінювання ризику вчинення рецидивних злочинів. Навчання моделі проведено на основі статистичних і динамічних характеристик та інформації про попередню кримінальну діяльність 13010 осіб, що відбувають покарання в установах виконання покарань України. Для виявлення значущих факторів, які корелюють зі схильністю засуджених до повторного вчинення злочинів, та ідентифікації ув'язнених з високим рівнем рецидивізму застосовано прогностні алгоритми машинного навчання. Визначено оптимальну модель, яка може бути використана для прогнозування ймовірності вчинення повторних злочинів для нових наборів даних. Запропонована модель може надати надійну інформаційно-правову підтримку при ухваленні судових

UDC 342.95

**MACHINE LEARNING MODELS  
FOR LEGAL INFORMATION  
SUPPORT COURT ACTIVITIES****Kovalchuk, Olga**

Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of the Department of Theory of Law and Constitutionalism

Western Ukrainian National University  
(Ternopil)

Artificial intelligence algorithms based on machine learning systems are already being used today to provide information and legal support for making effective judicial decisions in technologically advanced countries around the world. Ukraine has also started the digital transformation of the judicial system. Building a reliable information base for the courts requires the use of modern data science methods and innovative information technologies for high-quality analysis of large data volumes.

The article presents a scoring model for assessing the risk of recidivism. The model was trained based on statistical and dynamic characteristics and information about the previous criminal activities of 13,010 persons serving sentences in the penitentiary institutions of Ukraine. Predictive machine learning algorithms were applied to identify significant factors that correlate with the propensity of convicts to re-offend and to identify prisoners with a high level of recidivism. The optimal model that can be used to predict the likelihood of repeat offenses for new data sets has been determined. The proposed model can provide reliable information and legal support when making judicial decisions regarding the possible imposition of bail, parole, participation in probation or suspended

рішень щодо можливого призначення застави, умовно-дострокового звільнення, участі у пробації чи умовного засудження. Застосування передових наукових досягнень у діяльності судочинних органів є необхідним у сучасному смарт-світі, однак лише за умови належного інформаційно-правового регулювання, що забезпечить баланс між правами засуджених та безпекою суспільства.

Автор аналізує сучасні виклики, що стоять перед судовою системою у контексті обробки великого обсягу документів та даних. Розглядає можливість застосування методів машинного навчання для автоматизації процесів аналізу та обробки правової інформації, виявлення закономірностей та прийняття обґрунтованих рішень. Через приклади та аналіз реальних сценаріїв використання моделей машинного навчання у судочинстві, автор наводить переваги такого підходу та аналізує можливі виклики та обмеження.

Наведено приклади успішного впровадження моделей машинного навчання у судочинстві та акцентовано увагу на їхні переваги, такі як підвищена швидкість та точність прийняття рішень, а також покращення якості судочинства. Висновки статті підкреслюють важливість подальшого дослідження цієї теми та потенційного впровадження моделей машинного навчання в практику судочинства для підтримки суддів у їхній роботі.

Ключові слова: судова система, цифрова трансформація, інформаційно-правове забезпечення, інформаційні технології, машинне навчання

#### Постановка проблеми.

Сучасні інформаційні технології та дані, зібрані розумними пристроями, поступово трансформують наш світ у “розумне середовище”. Діджиталізація торкнулася і правосуддя, яке апріорі було суто “людською” сферою діяльності. Найближчим наука про дані (data science), штучний інтелект (artificial intelligence, ШІ), машинне навчання (machine learning) та технології великих даних (big data) стануть невід’ємними атрибутами інформаційно-аналітичної діяльності судів, створюючи нові можливості для ухвалення результативних судових рішень, позбавлених упередженості та суб’єктивності [1]. Ці технології можуть забезпечити органи судової влади дієвими інструментами для аналізу великих, складних масивів інформації у швидко мінливому онлайн-середовищі. Для підвищення ефективності й прозорості правосуддя судові органи мають орієнтуватися на інформаційні технології та

sentences. The application of advanced scientific achievements in the activities of the judiciary is necessary in the modern smart world, but only if proper information and legal regulation are in place that will ensure a balance between the rights of convicts and public safety.

The author analyses the current challenges faced by the judicial system in the context of processing large volumes of documents and data. He considers the possibility of applying machine learning methods to automate the processes of analysing and processing legal information, identifying patterns and making informed decisions. Through examples and analysis of real-life scenarios of using machine learning models in legal proceedings, the author outlines the advantages of this approach and analyses possible challenges and limitations.

Examples of successful implementation of machine learning models in judicial proceedings are provided and their advantages, such as increased speed and accuracy of decision-making, as well as improved quality of judicial proceedings, are highlighted. The conclusions of the article emphasise the importance of further research on this topic and the potential implementation of machine learning models in judicial practice to support judges in their work.

Keywords: judicial system, digital transformation, information and legal support, information technologies, machine learning

науку про дані, що надають інноваційні можливості для вдосконалення інформаційно-правової діяльності органів правосуддя [2, 3].

Інноваційні технології можуть зменшити суб'єктивність при ухваленні судових рішень і спростити доступ до правосуддя. У той же час при впровадженні інноваційних ІТ у практику судочинства необхідно враховувати існуючі небезпеки та ризики використання алгоритмів на основі систем штучного інтелекту, такі як дискримінація та упередження [4, 5, 6].

Інформаційно-правове забезпечення діяльності судів спрямоване на створення ефективних інструментів для застосування законодавства в судочинстві [7] і включає наступні складові:

- створення і використання електронних баз даних законодавства України, зокрема Єдиного державного реєстру судових рішень, Єдиного реєстру досудових розслідувань тощо;
- впровадження автоматизованих систем правової інформації для пошуку, систематизації й аналізу законодавства, судових рішень;
- забезпечення суддів та працівників апарату суду вільним безоплатним доступом до правових баз даних;
- створення спеціалізованих внутрішніх мереж для обміну правовою інформацією між судами різних інстанцій;
- підготовку суддівськими органами оглядів судової практики, узагальнень правових позицій судів;
- надання методичної допомоги суддям з питань застосування законодавства;
- вивчення та впровадження передового міжнародного досвіду організації судової діяльності;
- забезпечення відкритого доступу громадян до судових рішень через Інтернет.

Інформаційне забезпечення суду включає в себе створення та використання інформаційних систем, баз даних, мереж для збору, обробки, зберігання та поширення інформації, необхідної для функціонування судової системи. До інформаційних систем суду належать автоматизована система документообігу суду, автоматизована система діловодства суду, система відеофіксації судового процесу, Єдина судова інформаційно-телекомунікаційна система (ЄСІТС) [8] та інші. Інформаційне забезпечення сприяє прозорості та відкритості судової влади, доступу громадян до правосуддя та інформації про діяльність судів. Важливим є створення електронного суду, який дозволить подавати документи, вести справи, здійснювати відеотрансляції судових засідань онлайн. Питання інформатизації судів регулюється законами України «Про судоустрій і статус суддів» [9], «Про доступ до судових рішень» [10] та іншими нормативними актами. Впровадження інформаційних технологій значно підвищує ефективність роботи судової системи та спрощує доступ до правосуддя.

Вітчизняна система судочинства вже розпочала процес цифрової трансформації, однак поки досягла менших успіхів у впровадженні досягнень новітніх інноваційних технологій в інформаційно-аналітичне забезпечення органів суду порівняно з технологічно розвиненими країнами світу. Наукові дослідження з питань розробка прикладних моделей на основі науки про дані, штучного інтелекту та машинного навчання, які б могли надати інноваційні можливості для аналітично-інформаційної діяльності органів правосуддя та стати складовими ЄСІТС України, поки є вкрай рідкісними і потребують подальшого вдосконалення [11–13].

Аналіз останніх досліджень та публікацій. Проблеми використання сучасних інформаційних технологій і науки про дані для розширення інноваційних можливостей аналітично-інформаційної діяльності органів судочинства досліджували ряд науковців [14, 15]. Р. Зейрек та інші розглядали інтелектуальний аналіз даних і науку про дані як інновацію в юридичній сфері [16]. Р. Беркс досліджував можливість застосування статистичних процедур машинного навчання на основі великих даних для прогнозування повторного вчинення правопорушень особами, які вже перебувають відбувають покарання, за допомогою [17]. А. Бабута та інші аналізували переваги застосування алгоритмів машинного

навчання для підтримки ухвалення судових рішень, зокрема щодо можливості передбачення схильності осіб до майбутніх злочинів [18]. М. Гасемі та інші застосовували алгоритми машинного навчання для оцінювання ризиків та підтримки ухвалення оптимальних рішень у судочинстві [19]. Сьогодні все частіше застосовують науку про дані та комп'ютерне моделювання для підтримки ухвалення оптимальних судових рішень.

Метою роботи є побудова скорингової моделі для оцінювання ризику вчинення рецидивних злочинів на основі аналізу статистичних та динамічних даних ув'язнених та ідентифікації ув'язнених з високим рівнем рецидивізму. Отримані результати можуть надати органам судочинства релевантну інформацію для ухвалення судових рішень щодо можливого розміру застави, участі у пробації, умовного засудження, «умовно-дострокового звільнення від відбування покарання і заміну невідбутої частини покарання більш м'яким» [20].

Виклад основних результатів дослідження. Штучний інтелект вже проник у майже всі сфери нашого життя. Він може стати найбільш руйнівною технологією з часів промислової революції. Машинні прогнози не завжди є 100% точними. Вартість помилок різко зростає, коли ШІ застосовується у завданнях з високими ставками, таких як життя і свобода людей. Проте прогностичне правосуддя вже існує сьогодні. Алгоритми ШІ дедалі частіше використовують у судочинстві для прогнозування можливих злочинів та оцінки ймовірності рецидивів, у судах при призначенні застави чи покарання, встановленні вини чи невинуватості. Судові органи розвинених країн дедалі частіше застосовують у своїй аналітичній діяльності методи прогнозного картографування злочинів, передбачення можливих правопорушників і жертв злочину, прогнозування особи злочинця. Використання алгоритмів машинного навчання для інформаційної підтримки судових рішень може бути надзвичайно корисним для судової системи. Однією з найважливіших переваг ШІ є те, що його алгоритми можна ефективно застосовувати для зменшення кількості несправедливих вироків [21]. Одним з останніх успішних застосувань ШІ у судочинстві є штучна нейронна мережа, яка може ідентифікувати злочинців за специфічними рисами обличчя. Алгоритм продемонстрував 90% точності. Система здатна ідентифікувати особу на фото, знятих з різних ракурсів. Як і за інтонацією голосу, ШІ може знайти приховані емоції на фото з нейтральним виразом обличчя [22].

Судові органи дедалі частіше застосовують аналітику великих даних для стеження, розпізнавання образів та пошуку інформації у соцмережах. Ці технології використовують як інструмент, що допомагає органам правосуддя визначити засуджених з низьким ризиком рецидивізму для можливого дострокового звільнення. Алгоритми великих даних ґрунтуються на ймовірності вчинення злочину в майбутньому і призначені для інформування суддів про рекомендовану тривалість вироку. Однак існує велика небезпека отримання некоректних результатів через упередженість алгоритмів щодо певних верств населення [23]. У судочинстві набули поширення статистичні підходи та інформаційні технології для аналізу доказів та оцінки їх відповідності до розслідування і судового процесу. Технологічний огляд документів із застосуванням машинного навчання дозволяє значно підвищити ефективність аналізу доказової бази та обґрунтованість судових рішень [16]. Впровадження моделей на основі алгоритмів машинного навчання може сприяти ефективності інформаційно-правового забезпечення діяльності судів, однак лише за умов врахування усіх небезпек та ризиків та належного правового регулювання [7, 24].

Скорингова модель для оцінювання ризику вчинення рецидивних злочинів

Скорингова модель (scoring model) – це математична модель, яку використовують як інструмент для зваженого оцінювання можливих стратегій на основі стратифікації ризиків. Її отримують у результаті застосування інструментів статистичних методів до емпіричних даних з метою виявити фактори, які чинять суттєвий вплив на ймовірність настання аналізованої події. Стратегії оцінюються за стандартними критеріями, які базуються на співвідношенні втрат і вигод, а потім ранжуються за підсумковими балами [25].

У роботі запропоновано скорингову модель машинного навчання для виявлення факторів, які корелюють зі схильністю засуджених до повторного вчинення злочинів та

прогнозування ймовірності скоєння майбутніх рецидивів. Емпіричний аналіз проведено на основі унікального набору даних, який складають статистичні дані та інформація про попередню злочинну діяльність 13010 ув'язнених, які відбувають покарання в установах виконання покарань України. Дослідження такого роду проводять на державному рівні у високорозвинених країнах світу з високим рівнем злочинності. В Україні розробка прикладних моделей на основі ШІ та машинного навчання для інформаційно-аналітичного забезпечення судочинних органів поки не набула розвитку.

При проведенні емпіричних досліджень використано наступну інформацію про попередню злочинну діяльність та індивідуальні характеристики засуджених: наявність рецидивів, стать, вік, вік на момент першого ув'язнення, вік на момент першого ув'язнення / засудження до умовного покарання, стать, сімейний стан, рівень освіти, місце проживання, наявність умовно-дострокових звільнень, мотивація до звільнення, кількість засуджень до відбування покарання в установах виконання покарань, кількість умовних засуджень, кількість особливо важких злочинів, кількість важких злочинів, кількість злочинів середньої тяжкості, кількість злочинів невеликої тяжкості.

Для побудови комп'ютерної моделі використано, інструменти сучасної комп'ютерної аналітики та дослідження даних Statistica Data Miner. Для виявлення прихованих закономірностей в початковому наборі даних про ув'язнених визначено найбільш значущі для побудови моделі фактори, які класифікують засуджених на «схильних» і «несхильних» до рецидивів (Таблиця 1).

<b>Фактор</b>	<b>Значущість</b>
Кількість умовних засуджень	1,00
Кількість особливо тяжких злочинів	0,41
Вік на момент першого ув'язнення	0,39
Наявність умовно-дострокових звільнень	0,26
Вік на момент першого ув'язнення / засудження до умовного покарання	0,25
Стать	0,25
Рівень освіти	0,09
Сімейний стан	0,07
Тип зайнятості	0,04
Місце проживання	0,03
Вік	0,03
Мотивація до звільнення	0,02

*Таблиця 1. Значущість аналізованих факторів для прогнозування ризику скоєння рецидивних злочинів*

Найбільш вагомими факторами для передбачення ймовірного ризику скоєння повторних злочинів у майбутньому є наступні характеристики: кількість умовних засуджень, кількість особливо важких злочинів, вік на момент першого ув'язнення, наявність умовно-дострокових звільнень та вік на момент першого ув'язнення / засудження до умовного покарання та стать. Їх було використано для побудови прогнозової скорингової моделі.

Використання паралельно різних моделей машинного навчання дало можливість виявити приховані залежності та підтвердити попередні висновки. Для отримання оптимальної прогнозової моделі та перевірки надійності отриманих результатів застосовано наступні алгоритми машинного навчання): стандартні дерева класифікації з розгортанням, стандартна класифікація CHAID з розгортанням, підсилення дерев класифікації з розгортанням, інтелектуальний вирішувач проблем з розгортанням, метод опорних векторів з розгортанням (класифікація), MARSplines для класифікації з розгортанням (Рис. 1).

Проведено порівняння побудованих прогнозних моделей та обрано оптимальну. Для оцінки точності прогнозу застосовано перевірку на прикладах («hold-out» sample) та

перехресну перевірку (cross-validation). Оптимальна модель може правильно передбачити рецидивну приналежність з 98% точності. Відсоток правильного передбачення для категорії «несхильних» до рецидивів засуджених становить 98%.

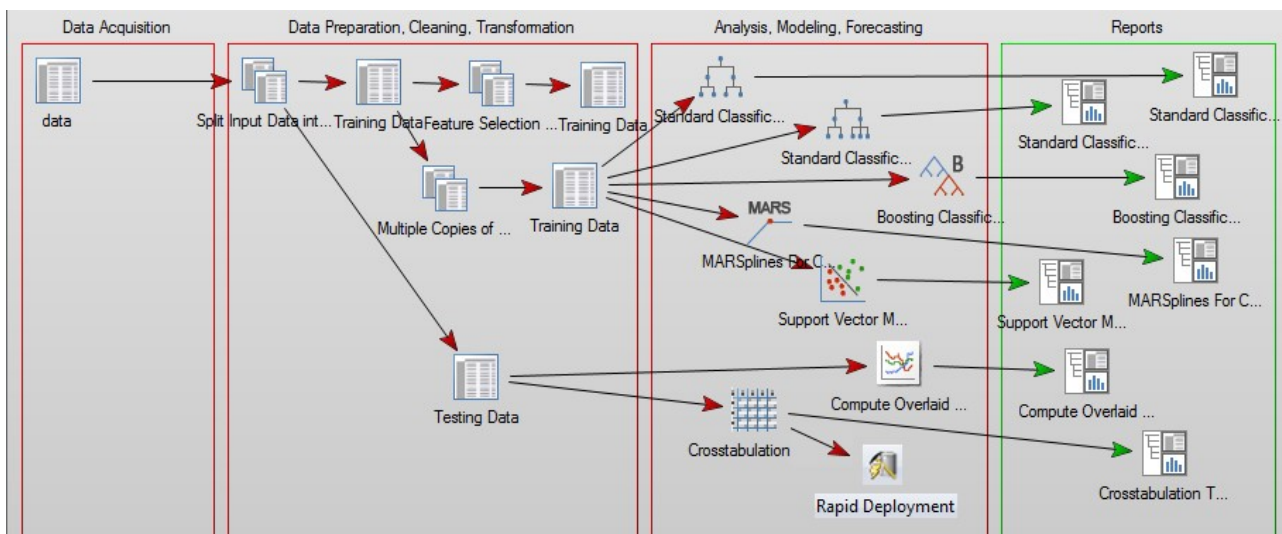


Рисунок 1. Робоча область Data Miner: етапи побудови скорингової моделі

Boosting Trees with Deployment є найкращою серед доступних моделей для прогнозування результатів. Моделі машинного навчання можуть бути застосовані для прогнозування ймовірності скоєння рецидивів для нових наборів даних та стати частиною інформаційно-правового забезпечення діяльності судочинної системи.

Висновки. У роботі побудовано скорингову модель для передбачення ймовірності вчинення кримінальних рецидивів особами, що вже були раніше засудженими до відбування покарання в установах виконання покарань, на основі їхніх індивідуальних статистичних і динамічних характеристик та інформації про попередню злочинну діяльність. Для визначення найбільш суттєвих факторів, що впливають на схильність засуджених до повторного вчинення злочинів та виявлення ув'язнених з високим рівнем рецидивізму застосовано різні прогнозні моделі на основі систем машинного навчання. Оптимальна прогнозна модель може бути застосована для нових наборів даних і надати надійну підтримку органам правосуддя при ухваленні ефективних рішень. Наприклад, при встановленні можливості призначення застави, умовного засудження чи умовно-дострокового звільнення.

Інформаційно-правова діяльність сучасних органів судочинства потребує ефективного використання даних, інструментів аналітики, ШІ та машинного навчання, вдосконалення доказової бази а також покращення методів аналізу та підвищення ефективності вимірювання та оцінювання. ШІ та машинне навчання ніколи не замінять суддів, однак можуть суттєво скоротити їхній час на виконання рутинних завдань, які не потребують високої кваліфікації, та надати надійну інформаційну підтримку при ухваленні судових рішень. Використання передових технологій в роботі судової системи є важливим в сучасному цифровому світі, проте це має відбуватися лише за умов належного правового регулювання. Таке регулювання має забезпечувати баланс між правами засуджених та безпекою суспільства при прийнятті судових рішень.

#### Посилання:

1. Greenstein S. Preserving the rule of law in the era of artificial intelligence (AI). *Artificial Intelligence and Law*. 2022. Vol. 30. pp. 291-323. DOI: 10.1007/s10506-021-09294-4.
2. Yassine S., Esghir M., Ibrihich O. Using Artificial Intelligence Tools in the Judicial Domain and the

- Evaluation of Their Impact on the Prediction of Judgments. *Procedia Computer Science*. 2023. Vol. 220. pp. 1021-1026. DOI: 10.1016/j.procs.2023.03.14.
3. Garat D., Wonsever D. Automatic Curation of Court Documents: Anonymizing Personal Data. *Information*. 2022. Vol. 13, article 27. DOI: 0.3390/info13010027.
4. Єдина судова інформаційно-телекомунікаційна система (ЄСІТС). Вища рада правосуддя. URL: <https://hcj.gov.ua/page/yesits>.
5. Про судоустрій і статус суддів : Закон України від 02.06.2016 р. URL : <https://zakon.rada.gov.ua/laws/show/1402-19#n1452>.
6. Теремецький В., Сімакова Н. Правові засади державного нагляду за здійсненням електронної роздрібної торгівлі лікарськими засобами. *Збірник наукових праць «Приватне право і підприємництво»*. 2023. № 23. С. 35-42. DOI: <https://doi.org/10.32849/2409-9201.2023.23.4>.
7. Про доступ до судових рішень : Закон України від 22.12.2005 р. URL : <https://zakon.rada.gov.ua/laws/show/3262-15#Text>.
8. Белов Д. М., Белова М. В. Штучний інтелект в судочинстві та судових рішеннях, потенціал та ризику. *Науковий вісник Ужгородського національного університету*. Серія: «Право». 2023. Том 2. № 78. С. 315-319. DOI: <https://doi.org/10.24144/2307-3322.2023.78.2.50>.
9. Куракін О., Скрябін О. Особливості правового регулювання використання штучного інтелекту в Україні. *Вісник Харківського національного університету імені В. Н. Каразіна*. Серія «Право». 2023. № 36. DOI: 10.26565/2075-1834-2023-36-04.
10. Припутень Д. Адміністративно-правове регулювання інформаційного забезпечення діяльності судів. *Науковий вісник Дніпропетровського державного університету внутрішніх справ*. 2023. № 1. С. 144-149. DOI: 10.31733/2078-3566-2023-1-144-149.
11. Berezka K., Kovalchuk O., Banakh S., Zlyvko S., Hrechaniuk R. A Binary Logistic Regression Model for Support Decision Making in Criminal Justice. *Folia Oeconomica Stetinensia*. 2022. Vol. 22 (1), pp. 1-17. DOI: 10.2478/fofi-2022-0001.
12. Kovalchuk O., Banakh S., Masonkova M., Burdin V., Zaverukha O., Ivanytsky R. A Scoring Model for Support Decision Making in Criminal Justice. *Proceedings of the 12th International Conference on Advanced Computer Information Technologies (Ruzomberok, Slovakia, 26–28 Sept. 2022)*, pp. 116-120. DOI: 10.1109/ACIT54803.2022.9913182.
13. Kovalchuk O., Banakh S., Masonkova M., Berezka K., Mokhun S., Fedchyshyn O. Text Mining for the Analysis of Legal Texts. *Proceedings of the 12th International Conference on Advanced Computer Information Technologies (Ruzomberok, Slovakia, 26–28 Sept. 2022)*, pp. 502-505. DOI: 10.1109/ACIT54803.2022.9913169.
14. Bhupatiraju S., Chen D. L., Joshi S. The Promise of Machine Learning for the Courts of India. *National Law School of India Review*. 2021. Vol. 3(2), article 10. URL : <https://repository.nls.ac.in/cgi/viewcontent.cgi?article=1439&context=nlisr>.
15. Saravanan P., Selvaprabu J., Raj L. A., Khan A., Sathick K. Survey on crime analysis and prediction using data mining and machine learning techniques. *Lecture Notes in Electrical Engineering*. 2021. Vol. 688, pp. 435-448. DOI: 10.1007/978-981-15-7241-8\_31.
16. Zeyrek R., Selvi S., Kalender M. F., Nesli S. Data science and criminal justice. *Digicrimus*. 06, Jun 2021. URL : <https://www.digicrimjus.com>.
17. Berks R. *Machine Learning Risk Assessments in Criminal Justice Settings*. Springer. 1st ed. 2019. 187 p.
18. Babuta A., Oswald M., Rinik C. *Machine Learning Algorithms and Police Decision-Making*. University of Winchester. 2018. 45 p.
19. Ghasemi M., Anvari D., Atapour M., Wormith J. S., Stockdale K. C., Spiteri R. J. The Application of Machine Learning to a General Risk–Need Assessment Instrument in the Prediction of Criminal Recidivism. *Criminal Justice and Behavior*. 2020. Vol. XX. No. X. pp. 1-21.
20. Про умовно-дострокове звільнення від відбування покарання і заміну невідбутої частини покарання більш м'яким: Постанова Пленуму Верховного суду України № 2 від 26.04.2002. URL : <https://zakon.rada.gov.ua/laws/show/v0002700-02#Text>.
21. Algorithms are being used to convict criminals and decide jail time. *World Economic Forum*. 2018. URL : <https://www.weforum.org>.
22. Fingerhood M. The Misuse of Big Data Algorithms in the United States Criminal Justice System. *Toward Data Science*, 26, Jun, 2019. URL : <https://towardsdatascience.com>.
23. Your next lawyer could be a machine. *World Economic Forum*. 2017. URL : <https://www.weforum.org>.
24. Теремецький В., Дуліба Є. Деякі питання надання вторинної безоплатної правової допомоги в адміністративному судочинстві. *Збірник наукових праць «Приватне право і підприємництво»*. 2023. № 23. С. 220-227. DOI: <https://doi.org/10.32849/2409-9201.2023.23.36>.

25. Jin F. Y. Interpretable Machine Learning Credit Scoring Model. Asia Pacific University. 2019. 55 p.

Статтю було подано	12.03.2024	The article was submitted
Статтю було доопрацьовано	15.03.2024	The article was revised
Статтю було прийнято	20.03.2024	The article was accepted