

УДК 004

JEL I21, I23, O33, C38, C55

DOI: 10.32782/2786-8273/2025-10-5

Гужва В.М.

кандидат економічних наук, доцент,
Київський національний економічний університет
імені Вадима Гетьмана
ORCID: <https://orcid.org/0000-0002-0379-1480>

Volodymyr Huzhva

Kyiv National Economic University
named after Vadym Hetman

ТЕХНОЛОГІЯ EDUCATIONAL DATA MINING ТА ЇЇ ВИКОРИСТАННЯ В ЦИФРОВІЙ ТРАНСФОРМАЦІЇ АКАДЕМІЧНИХ УСТАНОВ

EDUCATIONAL DATA MINING TECHNOLOGY AND ITS USE IN THE DIGITAL TRANSFORMATION OF ACADEMIC INSTITUTIONS

Анотація. У статті розглянуто технологію Educational Data Mining (EDM) як один із ключових інструментів цифрової трансформації академічних установ. Показано, що EDM забезпечує перехід від інтуїтивного прийняття рішень до дато-орієнтованого управління освітою. Розкрито основні методи та алгоритми (класифікація, кластеризація, асоціативні правила, аналіз соціальних мереж, видобування послідовностей), що застосовуються для прогнозування успішності, персоналізації навчальних траєкторій, оптимізації програм та підтримки стратегічного розвитку університетів. Проведено аналіз інструментів EDM – від вбудованих засобів LMS і спеціалізованих академічних систем до BI-платформ і мов програмування. Запропоновано модель багаторівневого застосування EDM: рівень викладача, факультету та університету. Обґрунтовано переваги EDM для підвищення якості освіти, а також окреслено виклики, пов'язані з цифровою зрілістю, стандартизацією та етичними аспектами роботи з освітніми даними.

Ключові слова: Educational Data Mining, аналітика навчання, цифрова трансформація, персоналізація освіти, вищі навчальні заклади, освітні дані, BI-платформи.

Abstract. Introduction. The rapid digitalization of higher education produces vast data flows in Learning Management Systems (LMS), where each student's activity leaves a trace. Yet the presence of data alone does not ensure effective decision-making. Many institutions still rely on intuition and fragmented observations, limiting opportunities for personalization and quality improvement. Educational Data Mining (EDM) transforms raw data into actionable knowledge, supporting a systematic and evidence-based approach to educational management. **Purpose.** The article explores EDM as a driver of digital transformation in academic institutions, emphasizing its essence, methodological foundations, software tools, and applications in teaching, learning, and governance. **Methods.** The study employs a systematic literature review, comparative analysis of EDM methods (classification, clustering, association rules, social network analysis, sequential pattern mining), and evaluation of software tools, ranging from LMS analytics and specialized research platforms to BI systems and programming languages. A practical case of integrating Microsoft Teams (Grades) with Power BI is also analyzed. **Results.** Findings show that EDM supports education at three complementary levels: (1) course level – early risk detection, personalized learning paths, quality improvement; (2) faculty level – performance monitoring, curriculum optimization, workload balancing; (3) institutional level – strategic decision-making, retention analysis, and accreditation support. EDM forms a holistic ecosystem of instruments that create a unified evidence base for academic management. **Conclusion.** EDM fosters a culture of data-driven education, where information becomes a strategic resource for sustainable institutional development and competitiveness. Key challenges include low digital maturity, lack of unified methodological standards, and ethical concerns regarding the use of student data. Future research should prioritize multi-algorithmic integration, development of universal models combining academic and non-academic factors, and closer alignment of EDM with decision-support systems in higher education.

Keywords: Educational Data Mining, learning analytics, digital transformation, personalization, higher education, academic performance, BI platforms.

Постановка проблеми. Сучасна епоха характеризується стрімкою цифровізацією усіх сфер суспільного життя, де дані перетворилися на один з найважливіших стратегічних ресурсів, що визначає розвиток економіки, науки та освіти. У так званій «економіці даних» вміння збирати, аналізувати та використовувати інформацію для прийняття обґрунтованих рішень стає ключовим конкурентним перевагою. Система освіти, як інтегральна частина суспільства, не лише реагує на цей глобальний тренд, але й активно його формує. Масовий перехід академічних установ на використання систем управління навчанням (LMS), таких як Moodle, Canvas, Blackboard та

інші, призвів до накопичення безпрецедентних масивів даних про навчальний процес. Кожна дія студента – перегляд лекції, виконання тесту, участь у форумі – залишає цифровий слід, створюючи багатий матеріал для аналізу.

Однак виникає проблема: наявність даних сама по собі ще не гарантує якісних результатів. Традиційно прийняття рішень в освіті (щодо індивідуальної траєкторії студента, ефективності курсу, розподілу ресурсів) часто ґрунтувалося на інтуїції, досвіді та обмежених спостереженнях. Цей підхід є неефективним у умовах масової освіти та не дозволяє в повній мірі реалізувати потенціал кожної людини. Існує гостра необхідність переходу від

суб'єктивних оцінок до дато-орієнтованого підходу, що дозволить підвищити якість освіти, її ефективність та забезпечити справжню персоналізацію навчання. Саме тут на допомогу приходить Educational Data Mining (EDM) – міждисциплінарна галузь знань, що поєднує педагогіку, інформатику та статистику.

Аналіз останніх досліджень і публікацій. В дослідженнях останніх років акцентується увага на використанні інтелектуального аналізу даних (EDM) та аналітики навчання (LA) у вищій освіті для підвищення якості освітнього процесу та прийняття рішень на основі даних. Так, М. Діас-Чоке та ін. [1] зазначають, що в умовах таких форс-мажорних обставин, як, наприклад, пандемія COVID-19, основним джерелом даних стали системи управління навчанням (LMS), а інтелектуальний аналіз даних може здебільшого застосовуватися для прогнозування успішності, зниження відсіву студентів та покращення якості навчання. Автори рекомендують інтегрувати EDM у політику академічного управління, щоб забезпечити сталість отриманих результатів при поверненні до очного навчання. Х. Аль-Доуа, Х. Аль-Самаррає та В.М. Фаузі [2] провели комплексний огляд 402 досліджень, що охоплювали чотири ключові напрями: комп'ютерна аналітика навчання (CSLA), прогнозна аналітика (CSPA), поведінкова аналітика (CSBA) та аналітика візуалізації (CSVA). Результати свідчать, що EDM і LA здатні підвищити індивідуалізацію навчання та надати університетам інструменти для постійного вдосконалення освітніх процесів. Р. Аларкон та ін. [3] розробили аналітичну систему для обробки академічних даних, яка включала вісім етапів – від збору та зберігання даних до візуалізації та підтримки прийняття рішень. Система тестувалася на даних SQL Server та файлах Excel приймальної комісії, використовуючи алгоритми опорних векторів (Support Vector Machines (SVM)), випадкових лісів (Random Forests (RF)), k найближчих сусідів (K-Nearest Neighbors (KNN)) та нейронних мереж, де найвища точність (94,21%) була досягнута методом SVM. Хітас та ін. [4] представили інтерактивну систему на основі EDM, яка аналізує онлайн-дані студентів до, під час та після COVID-19. Вона дозволяє прогнозувати прогрес студентів і коригувати навчальні траєкторії, роблячи освітній досвід більш персоналізованим та ефективним. А. Фернандес та ін. [5] дослідили ініціативи цифрової трансформації (DTI) у 39 університетах. Виявлено, що 24% DTI спрямовані на підвищення якості освіти, а серед технологій домінують розширена аналітика (23%), хмарні рішення (20%) та штучний інтелект (16%). Лише 25% ВНЗ мають цілісну цифрову стратегію, що свідчить про низький рівень цифрової зрілості сектору. К. Аулакха та ін. [6] підкреслюють роль EDM у розвитку електронного навчання, що стало ключовим після пандемії, а С. Ромеро та С. Вентура [7] систематизували методи EDM/LA та окреслили майбутні тенденції розвитку цієї галузі, включно з використанням великих даних та академічної аналітики. Ф. Гарсія та ін. [8] визначили необхідність оцінки цифрової зрілості ЗВО перед впровадженням цифрової трансформації, що охоплює інфраструктуру, цифрову компетентність викладачів і студентів та методологічну інтеграцію технологій. Й.Х. Крісанто та Г. Абдулла [9] довели ефективність поєднання різних методів EDM – класифікації (ID3, SVM), кластеризації (k-Means, k-Medoids), правил асоціації (Apriori) та виявлення аномалій (DBSCAN) – для стратегічного управління академічною успішністю.

На основі проведеного аналізу публікацій можна виділити наступні невирішені проблеми: а) недостатня інтеграція EDM/LA у стратегічні процеси управління університетами; б) обмежене використання мультиалгоритмічного підходу для аналізу освітніх даних; в) низький рівень цифрової зрілості, що гальмує впровадження комплексних DTI; г) відсутність єдиних стандартів застосування EDM у гібридних і постпандемічних моделях навчання; д) потреба у розробці універсальних моделей аналізу, які враховують як академічні, так і неакадемічні фактори успішності студентів.

Метою даної статті є комплексне дослідження технології Educational Data Mining як інструменту цифрової трансформації освіти. У роботі планується розкрити сутність та основні методи EDM, провести огляд сучасного програмного інструментарію та проаналізувати практичні аспекти її застосування в академічних установах. Особлива увага буде приділена тому, як застосування EDM може сприяти підвищенню академічної успішності студентів через механізми раннього попередження ризиків, персоналізації навчального контенту та оптимізації навчального процесу в цілому. Стаття спрямована на те, щоб продемонструвати практичну цінність EDM для викладачів, адміністраторів та керівників освітніх установ.

Виклад основного матеріалу дослідження. Educational Data Mining (EDM) – це міждисциплінарна наукова галузь, що знаходиться на стику педагогіки, інформатики, статистики та машинного навчання. Її основним завданням є розробка, адаптація та застосування методів і алгоритмів для виявлення шаблонів і закономірностей у великих масивах даних, які генеруються в освітньому середовищі. EDM не просто збирає дані, а трансформує їх у глибоке розуміння навчальних процесів, що дозволяє перейти від описових констатацій до предиктивного аналізу та персоналізованих освітніх траєкторій. Це процес «видобутку знань» (knowledge discovery) з «сирих» освітніх даних.

EDM як окрема дисципліна сформувалася на початку 2000-х років, але її витоки лежать у застосуванні методів інтелектуального аналізу даних (Data Mining) в освіті ще в 1990-х. Аналіз публікацій [10–12] свідчить про те, що в історії виникнення та становлення Educational Data Mining можна виділити наступні події:

1) 1995–2005 рр.: *Передісторія*. Перші дослідження були зосереджені на аналізі лог-файлів навчальних середовищ та створенні моделей знань. Цей період характеризується розрізненими, ізольованими дослідженнями.

2) 2005 р.: *Зародження як окремої галузі*. Перша міжнародна конференція з Educational Data Mining (EDM) стала ключовою віхою. Вона об'єднала дослідників з різних дисциплін і формально закріпила статус EDM як самостійного наукового напрямку. Ця конференція щорічно проводиться і донині.

3) 2008 р.: *Формування академічного ядра*. Було засновано Міжнародне товариство з Educational Data Mining та запущено Міжнародний журнал Educational Data Mining (IJEDM). Це створило централізовану платформу для публікації результатів досліджень, обміну ідеями та розвитку спільноти.

4) 2010–2020-ті рр.: *Інтеграція, масштабування та поглиблений аналіз*.

– *тісне переплетення з Learning Analytics (LA)*. Ці дві галузі розвиваються паралельно, взаємно доповнюючи одна одну;

– *інтеграція в комерційні продукти*. Великі вендори LMS (Moodle, Canvas, Blackboard) почали впроваджувати інструменти аналітики та предиктивні моделі прямо в свої платформи;

– *доступність інструментів*. З'явилися зручні платформи з графічним інтерфейсом (RapidMiner, KNIME тощо), що зробили EDM доступнішою для дослідників без глибоких знань програмування.

– *акцент на етиці та конфіденційності*. Зростання обсягів даних викликало активну дискусію про етичні норми, прозорість алгоритмів та захист персональних даних студентів.

Таким чином, EDM пройшла шлях від окремих академічних експериментів до потужного інструментарію, що активно впроваджується в освітній процес по всьому світу для його глибокої трансформації.

Часто поняття EDM та Learning Analytics (LA) використовуються як взаємозамінні, проте між ними існує тонка, але важлива різниця в акцентах [13]:

– *Educational Data Mining (EDM)* зосереджена на технічних аспектах та розробці нових алгоритмів для автоматичного дослідження даних. Її ядро становлять методи машинного навчання, статистичного аналізу та pattern recognition. EDM відповідає на питання «Які нові методи ми можемо застосувати, щоб знайти приховані закономірності в освітніх даних?».

– *Learning Analytics (LA)* більш орієнтована на практичне застосування отриманих інсайтів для поліпшення навчання та освітнього середовища. LA інтегрує результати EDM з педагогічною теорією, соціологією та когнітивними науками. Вона відповідає на питання «Як ми можемо використати ці закономірності, щоб покращити навчальний процес, підвищити успішність студентів та оптимізувати роботу установи?». Зв'язок між ними можна представити наступним чином: EDM є «двигуном» (engine), який забезпечує технологічну основу для аналізу, тоді як LA – це «автомобіль» (car), який використовує цей двигун для досягнення конкретних практичних цілей. Ці дві галузі тісно переплетені і разом формують потужний інструментарій сучасної дато-орієнтованої освіти.

В контексті вище зазначеного доцільним може бути визначення предмету та об'єкту дослідження EDM:

1) Об'єкт дослідження EDM – це освітні дані, що генеруються в процесі навчання. Ці дані надзвичайно різноманітні та структуровані:

а) дані про активність в LMS: логи входів, перегляд відеолекцій, скачування матеріалів, участь у форумах, виконання тестів та завдань;

б) оцінки та результати: бали за тести, завдання, курсові роботи, рейтинги, підсумкові оцінки;

в) демографічні дані: вік, стать, курс, спеціальність, географічне розташування, попередня освіта;

г) дані соціальних взаємодій: мережі зв'язків між студентами на форумах (хто кому відповідає), коментарі, спільна робота над проектами;

д) метадані про контент: тип навчального матеріалу (текст, відео, інтерактивний модуль), складність, прив'язка до теми.

2) Предмет дослідження EDM – це закономірності, шаблони та знання, які виявляються в цих даних і які описують навчальний процес. Зокрема, EDM вивчає:

а) моделі поведінки студентів: стилі навчання, стратегії вирішення задач, паттерни прокрастинації (типові шаблони поведінки);

б) чинники академічної успішності: які показники найкраще прогнозують результат;

в) ефективність педагогічних методів і навчальних матеріалів;

г) динаміку формування соціальних зв'язків у навчальних спільнотах.

Таким чином, EDM перетворює дані з пасивного продукту діяльності на активний інструмент для глибшого розуміння того, як відбувається навчання, і що можна зробити, щоб зробити його ефективнішим для кожної людини.

Методи EDM можна структуровано класифікувати за типами завдань, які вони вирішують, та видами аналізу, які вони дозволяють провести [14–17]. Кожен метод має свою специфіку та область найефективнішого застосування в освітньому контексті (табл. 1).

Таблиця 1

Методи Educational Data Mining

№	Назва методу	Суть методу	Застосування в EDM	Популярні алгоритми
1	2	3	4	5
1	Кластеризація (Clustering)	Це метод неконтрольованого навчання (unsupervised learning), метою якого є автоматичне розбиття набору даних на групи (кластери) таким чином, щоб об'єкти всередині однієї групи були максимально схожими між собою, а об'єкти з різних груп – максимально відмінними.	1) Виявлення типів студентів: розподіл студентів на групи за стилями навчання, моделями поведінки чи рівнем знань. Наприклад: – Кластер 1: «Відстаючі» (низька активність, рідкісні відвідування, погані оцінки). – Кластер 2: «Активні самостійники» (висока активність, перегляд всіх матеріалів, але низька участь у форумах). – Кластер 3: «Соціальні лідери» (висока активність на форумах, допомога іншим). 2) Сегментація навчального контенту: групування курсів або завдань за складністю або типом.	k-середніх (k-means), ієрархічна кластеризація, DBSCAN.

Продовження таблиці 1

1	2	3	4	5
2	Класифікація та прогностичне моделювання (Classification & Predictive Modelling)	Це метод контрольованого навчання (supervised learning). Модель навчається на історичних даних, для яких вже відомо правильне значення цільової змінної (наприклад, «склав»/«не склав»), а потім використовує набуті знання для прогнозування результату для нових даних.	1) Раннє виявлення студентів з ризиком невдачі: Це найпоширеніше застосування. Модель класифікує студентів за категоріями ризику (наприклад, «низький», «середній», «високий» ризик відчислення). 2) Прогнозування підсумкової оцінки. 3) Класифікація типів помилок, які припускаються студенти.	1) Дерева рішень (Decision Trees): зручні тим, що результати легко інтерпретувати («якщо студент не здав перше завдання і не заходив у систему 10 днів, то ризик = високий») 2) Наївний байєсівський класифікатор (Naive Bayes): часто використовується для класифікації текстів (наприклад, на форумах). 3) Метод опорних векторів (Support Vector Machines – SVM): ефективний для бінарної класифікації (наприклад, «склав»/«не склав») 4) Логістична регресія (Logistic Regression).
3	Асоціативне правило (Association Rule Mining)	Метод виявляє приховані зв'язки, залежності та асоціації між різними подіями або об'єктами у великих наборах даних. Результатом є правила типу «Якщо X, то Y» (if-then).	1) Аналіз ефективності навчальних матеріалів: виявлення зв'язків між використанням різних ресурсів і успішністю. Наприклад: {Перегляд відеолекції, Виконання інтерактивного тесту} → {Висока оцінка на іспиті} (певність 85%). 2) Побудова рекомендаційних систем: «Студенти, які переглядали цей матеріал, також переглядали ось це».	Apriori, FP-Growth.
4	Аналіз соціальних мереж (Social Network Analysis – SNA)	Метод аналізує структуру взаємин між учасниками соціальної мережі. В освітньому контексті мережа будується на основі взаємодій між студентами та викладачами (наприклад, хто кому відповідає на форумі).	1) Виявлення лідерів думок та ізольованих студентів. 2) Оцінка колаборації та групової роботи. 3) Визначення моделей поширення інформації в навчальній спільноті. 4) Прогнозування успішності на основі соціального статусу студента в мережі (наявність соціальних зв'язків часто корелює з успішністю).	Основні метрики: а) центральність (degree centrality), б) посередництво (betweenness centrality), в) щільність мережі.
5	Видобування послідовностей (Sequential Pattern Mining)	Метод виявляє типові послідовності подій або дій, які відбуваються в певному порядку. Це дозволяє аналізувати не просто наявність ознак, а їхню часову послідовність.	1) Аналіз навчальних шляхів (Learning Pathways): виявлення найефективніших послідовностей вивчення тем. Наприклад: «Студенти, які спочатку переглядають конспект, а потім роблять практичне завдання, досягають кращих результатів, ніж ті, хто робить навпаки». 2) Діагностика труднощів: визначення послідовностей дій, які зазвичай ведуть до неправильної відповіді. 3) Моделювання поведінки студента протягом усього курсу.	1) GSP (Generalized Sequential Pattern): один з перших алгоритмів, який використовує метод «рівень за рівнем» для видобутку послідовних шаблонів. 2) PrefixSpan (Prefix-projected Sequential pattern mining): використовує техніку проєкції префікса, що дозволяє зменшити обсяг даних, що обробляються на кожному кроці. 3) SPADE (Sequential Pattern Discovery using Equivalence classes): використовує структуру даних, що базується на еквівалентності, для швидкого видобутку послідовних шаблонів. 4) Clospan: модифікація алгоритму PrefixSpan, що фокусується на ефективному видобутку послідовних шаблонів у великих наборах даних. 5) BIDE (Bi-Directional Extension): алгоритм, що дозволяє знаходити довгі послідовності, зменшуючи кількість повторних обчислень через двостороннє розширення. 6) SPAM (Sequential pattern mining by pattern growth): використовує метод зростання шаблонів, щоб знаходити послідовності, не потребуючи генерації кандидатів.

Джерело: складено автором на основі [14–17]

На основі наведеної в таблиці інформації слід зазначити, що вибір конкретного методу або алгоритму в EDM безпосередньо залежить від поставленої педагогічної або дослідницької задачі. Найчастіше ці методи використовуються не ізольовано, а в комбінації, що дозволяє отримати найбільш повну та об'єктивну картину освітнього процесу. Наприклад, спочатку метод кластеризації може сегментувати студентів на групи, а потім для кожної групи можуть бути побудовані свої прогностичні моделі або виявлені свої асоціативні правила.

Успішне впровадження Educational Data Mining вимагає використання спеціалізованого програмного забезпечення. Інструменти можна класифікувати за рівнем складності, цільовою аудиторією та гнучкістю.

Станом на поточний момент можна виділити наступні категорії інструментів:

1) **Вбудовані інструменти LMS** (для викладачів та адміністраторів) – ці інструменти інтегровані безпосередньо в середовище системи управління навчанням, що робить їх найбільш доступними:

а) Аналітика в Moodle: набір базових звітів про активність студентів (перегляд матеріалів, участь у форумах, результати тестів). З появою Moodle Learning Analytics API з'явилася можливість створювати та використовувати прогностичні моделі (наприклад, модель «Студенти з ризиком не завершити курс») безпосередньо в LMS;

б) Canvas Analytics: пропонує більш просунуті та візуалізовані звіти порівняно з Moodle. Ключова особливість – Canvas Analytics for Predict – механізм раннього попередження, який присвоює студентам «рахунок ризику» на основі їхньої активності;

в) Blackboard Analytics: потужний, але дорогий комерційний рішення, орієнтований на аналітику на рівні всієї установи. Інтегрує дані з LMS, студентської інформаційної системи (SIS) та фінансових систем для стратегічного аналізу.

2) **Спеціалізовані академічні інструменти (для дослідників)** – створені для вирішення конкретних наукових завдань, часто мають вузьку спрямованість.

а) Kepler (SNAPP – Social Networks Adapting Pedagogical Practice): веб-інструмент для візуалізації та аналізу соціальних мереж на освітніх форумах. Дозволяє бачити лідерів думок, ізольованих студентів і структуру взаємодій. Простий у використанні, але працює лише з асинхронними форумами;

б) LOCO-Analyst (Learning Object Context Ontology Analyst): інструмент для аналізу контексту навчальної діяльності. Генерує звіти про якість навчальних матеріалів і соціальну взаємодію, прив'язуючи активність студентів до конкретних навчальних об'єктів. Складніший у налаштуванні, ніж Kepler.

3) **Платформи для предиктивної аналітики та інтервенцій (для викладачів та адміністраторів):** інструменти, орієнтовані на дію, а не лише на аналіз. Прикладом може слугувати OnTask (Home | OnTask) – потужна платформа з відкритим кодом для масованої персоналізованої комунікації. Імпортує дані з LMS, дозволяє налаштувати правила (наприклад, «IF оцінка < 50 THEN...») і автоматично відправляти студентам персоналізовані емейли з зворотним зв'язком, порадами чи ресурсами. Націлена на оперативне педагогічне втручання.

4) **Універсальні інструменти Data Science (для аналітиків та дослідників)** – потужні платформи з графіч-

ним інтерфейсом, що дозволяють проводити складний аналіз без глибокого програмування:

а) RapidMiner, KNIME: платформи для візуального програмування аналітичних процесів (data pipelines). Користувач перетягує блоки (завантаження даних, очищення, модель, візуалізація), що робить їх дуже популярними для швидкого прототипування моделей EDM;

б) WEKA (Waikato Environment for Knowledge Analysis): класичний набір програм з відкритим кодом, що містить величезну колекцію алгоритмів машинного навчання. Менш сучасний інтерфейс, але дуже потужний для навчання і експериментів;

в) Orange: інструмент з дуже інтуїтивним та візуально привабливим інтерфейсом. Ідеально підходить для ознайомлення з EDM та швидкої побудови інтерактивних дашбордів.

5) **Мови програмування (для data scientists та дослідників)** – надають максимальну гнучкість і контроль для глибокого кастомного аналізу:

– Python (з бібліотеками Pandas, Scikit-learn, NumPy, Matplotlib/Seaborn): де-факто стандарт у галузі Data Science. Pandas використовується для маніпуляції даними, Scikit-learn – для побудови машинного навчання, а бібліотеки для візуалізації – для створення звітів;

– R: особливо популярна в академічному середовищі для статистичного аналізу та створення публікаційно-якісних графіків. Має велику кількість спеціалізованих пакетів для різних видів аналізу.

6) **Інтегровані BI-платформи (для аналітиків, адміністраторів та керівництва)** – ця категорія включає потужні корпоративні інструменти, призначені для консолідації даних з різних джерел, побудови комплексних дашбордів та проведення глибокого аналізу для прийняття стратегічних рішень.

Такі платформи, як Microsoft Power BI, Tableau, Qlik Sense, є «важкою артилерією» аналітики. Вони можуть інтегрувати дані практично з будь-яких джерел: LMS (Moodle, Canvas), SIS, опитувань, фінансових систем та навіть зовнішніх баз даних. Ключова перевага: можливість створювати «єдину точку правди» (single source of truth) для всієї установи, зв'язуючи дані про навчання з адміністративними та фінансовими показниками. Вони ідеальні для: стратегічного аналізу на рівні університету або факультету – аналіз відтворюваності (retention rates), успішності випускників, ефективності програм, розподілу ресурсів. Відмінність від інших категорій полягає в наступному: вони менше зосереджені на педагогічних інтервенціях в реальному часі (як OnTask) і більше – на макротрендах і KPI для керівництва.

В табл. 2 приведено порівняльний аналіз описаних вище інструментів Educational Data Mining.

На основі даних можна зробити такий висновок: вибір інструменту повністю залежить від цілей та аудиторії, а саме:

1) для викладача, який хоче швидко побачити активність студентів у своєму курсі, достатньо вбудованих звітів LMS;

2) для дослідника, який вивчає динаміку спілкування на форумі, ідеально підійде Kepler;

3) для академічного декана, якому потрібно зрозуміти тенденції успішності на всіх курсах факультету та прогнозувати вибуття, необхідна BI-платформа (наприклад, Power BI), яка інтегрує дані з усіх курсів;

4) для команди підтримки студентів, яка хоче автоматизувати відправку нагадувань про консультації, найкращим рішенням буде OnTask;

5) для data scientist університету, який будує складну прогностичну модель з використанням AI, основним інструментом будуть Python/R.

Ідеальний сценарій для великої освітньої установи – це комбінація цих інструментів:

– Power BI/Tableau – для стратегічного аналізу на рівні керівництва;

– OnTask – для тактичних педагогічних інтервенцій на рівні курсів;

– Вбудована аналітика LMS – для щоденної роботи викладачів;

– Python/R – для наукових досліджень і розробки нових моделей.

Таким чином, інструменти EDM можуть утворювати екосистему, що забезпечує підтримку прийняття рішень на всіх рівнях освітньої установи.

Застосування Educational Data Mining має каскадний ефект, поширюючись від індивідуальної роботи зі студентом до стратегічного планування розвитку університету. Його реалізація відбувається на трьох ключових рівнях (табл. 3).

Таблиця 2

Підсумкова таблиця порівняння категорій інструментів EDM

Категорія	Основна аудиторія	Переваги	Недоліки	Ключові інструменти
1. Вбудовані в LMS	Викладачі, адміністратори	Простота, доступність, немає потреби в інтеграції	Обмежений функціонал, закрита архітектура	Moodle Analytics, Canvas Analytics, Blackboard Analytics
2. Спеціалізовані академічні	Дослідники	Глибина аналізу в конкретній галузі (наприклад, соціальні мережі)	Вузкоспеціалізовані, часто потребують налаштування	Kepler (SNAPP), LOCO-Analyst
3. Платформи інтервенцій	Викладачі, адміністратори	Орієнтація на дію, автоматизація підтримки	Потрібна якісна настройка правил	OnTask
4. Універсальні DS-інструменти	Аналітики, дослідники	Візуальне моделювання, широкий спектр алгоритмів	Потрібна експертиза в Data Mining	RapidMiner, KNIME, WEKA, Orange
5. Мови програмування	Data Scientists, дослідники	Максимальна гнучкість і контроль	Висока складність вивчення та використання	Python (Pandas, Scikit-learn), R
6. Інтегровані BI-платформи	Керівництво, адміністратори, аналітики	Інтеграція даних, потужна візуалізація, стратегічний аналіз	Вартість, складність розгортання	Microsoft Power BI, Tableau, Qlik Sense

Джерело: складено автором

Таблиця 3

Застосування технології EDM на різних рівнях управління академічною установою

Рівень управління	Управлінська ситуація	Як працює	Результат	Конкретні кейси використання
1	2	3	4	5
Рівень викладача/курсу: Оперативне втручання та оптимізація (На цьому рівні EDM виступає інструментом «педагогічного навігатора», що надає викладачу дані для прийняття обґрунтованих рішень щодо кожного студента та курсу в цілому).	1) Раннє виявлення студентів з ризиком неуспішності	Система (наприклад, вбудована в LMS або зовнішня, як OnTask) аналізує поведінкові дані: відвідуваність, активність у форумах, своєчасність здачі завдань, результати перших тестів. Алгоритми класифікації (дерева рішень, логістична регресія) присвоюють кожному студенту ймовірність неуспішного завершення курсу.	Викладач отримує автоматичне сповіщення або список студентів із зазначенням рівня ризику («червоний», «жовтий», «зелений» індикатор) ще до моменту, коли ситуація стане критичною.	Кейс 1 (Рівень викладача): Викладач інформатики в LMS бачить, що 5 студентів отримали «червоний» індикатор ризику після першого тесту. Система вказує, що основна причина – нульова активність на форумі. Викладач через OnTask відправляє цій групі автоматичне персоналізоване повідомлення: «Привіт! Бачу, що результат тесту може бути кращим. Запрошую тебе приєднатися до обговорення на форумі [посилання] – це допоможе краще засвоїти тему. Якщо виникли труднощі, завітай на консультацію у середу о 15:00».
	2) Персоналізація навчання та автоматизовані інтервенції	На основі аналізу попередніх успіхів і невдач система може: а) Автоматично адаптувати складність завдань (адаптивне навчання). б) Формувати персоналізовані рекомендації («вам корисним буде переглянути це відео»); в) Генерувати та відправляти автоматичні повідомлення через email або LMS. Наприклад, якщо студент не здав завдання, система може надіслати йому лист із заохоченням та посиланням на додаткові матеріали.	Підвищення мотивації та залученості студентів через індивідуальний підхід без значного збільшення навантаження на викладача.	

Продовження таблиці 3

1	2	3	4	5
Рівень викладача/курсу: Оперативне втручання та оптимізація (На цьому рівні EDM виступає інструментом «педагогічного навігатора», що надає викладачу дані для прийняття обґрунтованих рішень щодо кожного студента та курсу в цілому).	3) Аналіз якості навчального контенту та його вдосконалення	Аналізуються дані про взаємодію з матеріалами: «теплові карти» відеолекцій (які фрагменти переглядали кілька разів або пропускали), статистика використання текстів, презентацій, результатів тестів, прив'язаних до конкретних тем.	Викладач отримує об'єктивний звіт про те, які матеріали ефективні, а які – ні. Наприклад, якщо більшість студентів пропускає один і той же фрагмент лекції, це сигнал до її перезапису чи пояснення.	Кейс 1 (Рівень викладача): Викладач інформатики в LMS бачить, що 5 студентів отримали «червоний» індикатор ризику після першого тесту. Система вказує, що основна причина – нульова активність на форумі. Викладач через OnTask відправляє цій групі автоматичне персоналізоване повідомлення: «Привіт! Бачу, що результат тесту може бути кращим. Запрошую тебе приєднатися до обговорення на форумі [посилання] – це допоможе краще засвоїти тему. Якщо виникли труднощі, завітай на консультацію у середу о 15:00».
	4) Формування збалансованих груп для проєктів	Алгоритми кластеризації аналізують набір характеристик студентів (академічна успішність, соціальна активність, стиль навчання) і формують гетерогенні групи, де студенти з різними якістьями доповнюють один одного.	Підвищення ефективності групової роботи, зменшення конфліктів і забезпечення справедливого розподілу ролей.	
Рівень факультету/деканату: Тактичне управління якістю освіти (На цьому рівні EDM використовується для аналізу та оптимізації навчальних програм і процесів у межах факультету або деканату)	1) Аналіз успішності за спеціальностями та курсами	Агрегуються дані з усіх курсів спеціальності. Аналітики будують дашборди, які показують динаміку успішності, відсоток невдач по конкретним дисциплінам, середній бал.	Деканат може виявити «вузькі місця» в навчальному плані. Наприклад, якщо курс «Вища математика II» consistently має аномально високий відсоток неспад, це привід для перегляду його програми, зміни викладача або введення додаткової підтримки.	Кейс 2 (Рівень факультету): Аналітик деканату факультету комп'ютерних наук в Power BI виявляє, що курс «Програмування» має 40% неспад, причому найнижчі результати у студентів, які раніше отримали оцінку нижче В за курс «Алгоритмізація». Деканат приймає рішення зробити «Алгоритмізація» обов'язковою передумовою для запису на «Програмування».
	2) Моніторинг ефективності викладачів та навчальних програм	Аналізуються не лише результати студентів, а й дані про методи викладання: використання інтерактивних інструментів, активність на форумах, швидкість перевірки завдань, відгуки студентів.	Можливість об'єктивно оцінити внесок викладача, визначити найефективніші педагогічні практики та масштабувати їх, а також виявити потребу у підвищенні кваліфікації ППС.	
	3) Управління навчальним навантаженням	Система аналізує розподіл кредитів ECTS, складність курсів (на основі даних про успішність) та навантаження викладачів.	Оптимальний розклад, справедливий розподіл навантаження між викладачами та запобігання «завантаження» студентів надто складними курсами в один семестр.	

Продовження таблиці 3

1	2	3	4	5
Рівень університету/керівництва: Стратегічний аналіз і розвиток (На цьому рівні EDM інтегрується з даними з інших систем (SIS, CRM, фінансової) для підтримки стратегічних рішень керівництва)	1) Стратегічний аналіз набору, відрахування та випуску студентів	Побудова прогностичних моделей, що визначають фактори, які впливають на рішення абітурієнта вступити, а студента – продовжити навчання чи відчислитися. Аналізуються дані про вступні бали, географію, соціально-економічний статус, успішність.	Цільовий набір, розробка ефективних програм підтримки та стимулів для утримання студентів (retention rate), що безпосередньо впливає на фінансову стабільність університету.	Кейс 3 (Рівень університету): Кураторський відділ університету, використовуючи прогностичну модель, визначає, що студенти-першокурсники з інших міст, які проживають у гуртожитку і мають низький вступний бал, мають найвищий ризик відчислення в першому семестрі. Університет запускає для цієї цільової групи програму менторства та додаткових соціально-адаптаційних заходів, що дозволяє знизити відсоток відрахування на 15%.
	2) Прийняття управлінських рішень на основі даних	Аналіз ефективності різних підрозділів, популярності спеціальностей, витрат на навчальні програми.	Обґрунтовані рішення про відкриття чи закриття спеціальностей, розподіл бюджету, інвестиції в нові освітні технології чи інфраструктуру.	
	3) Підготовка звітів для акредитації	Автоматизоване формування звітів із показниками ефективності діяльності університету (кількість випускників, їхня працевлаштованість, середній бал, академічна успішність).	Значно спрощується процес проходження акредитації, всі необхідні дані є в єдиній системі та об'єктивно підтверджені.	

Джерело: складено автором

Практичний кейс. Прикладом реалізації технології EDM може слугувати використання інтегрованої аналітичної платформи Power BI для аналізу успішності студентів на основі індивідуальних результатів, які збираються за допомогою функції Grades у команді кожної групи в модулі Teams (рис. 1). У прикладі в якості вхідних даних використовуються індивідуальні результати (оцінки за виконання і захист різних видів завдань) в рамках опанування дисципліни «Системи управління знаннями», яку автор викладає в університеті (рис. 2). Результати проведеного EDM-аналізу за допомогою платформи Power BI наведені на рис. 3, а-в. Ці результати можуть бути використані викладачем для удосконалення навчально-методичних матеріалів та визначення потреби в індивідуальних консультаціях для окремих студентів.

Висновки. Технологія Educational Data Mining (EDM) довела свою ефективність як інструмент цифрової трансформації академічних установ, оскільки переводить управління освітою від інтуїтивних рішень до підходу, заснованого на даних. Вона забезпечує персоналізацію навчання, раннє виявлення ризиків, оптимізацію освітніх програм і стратегічний розвиток університетів. Аналіз показав, що EDM працює на різних рівнях: для викладачів – індивідуалізація навчальних траєкторій і вдосконалення матеріалів; для факультетів – підвищення ефективності програм та управління якістю; для керівництва університету – стратегічне планування й конкурентоспроможність. Її практична цінність полягає у використанні екосистеми інструментів (LMS-аналітика, BI-платформи, Python/R тощо), що охоплю-

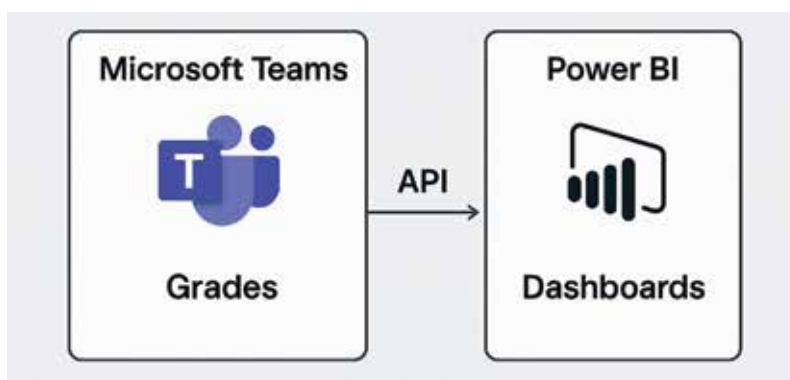


Рис. 1. Спрощена схема взаємодії інтегрованої аналітичної Power BI з модулем Teams за допомогою API (Application Programming Interface)

Джерело: розроблено автором

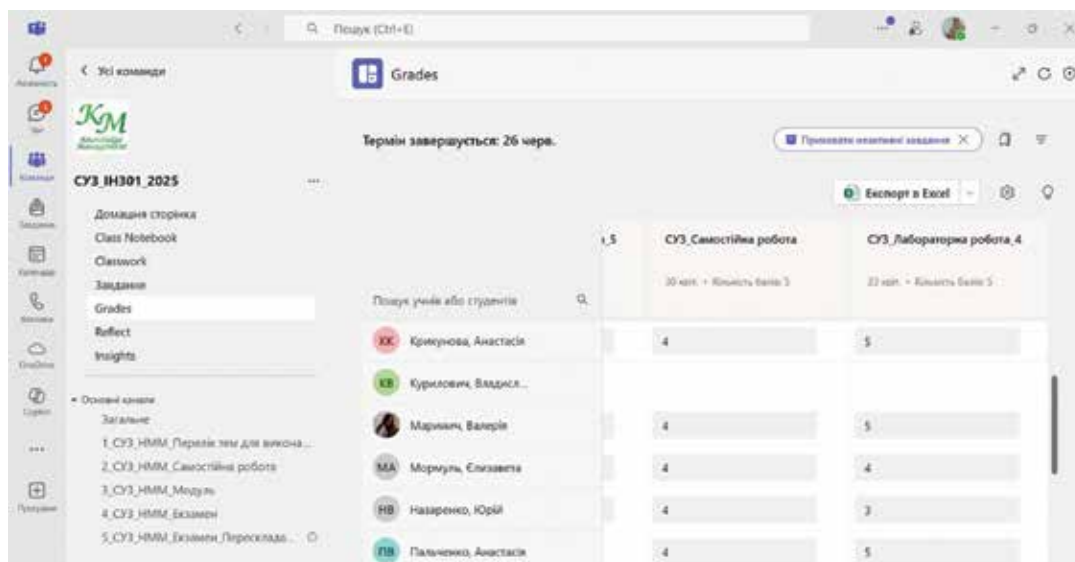
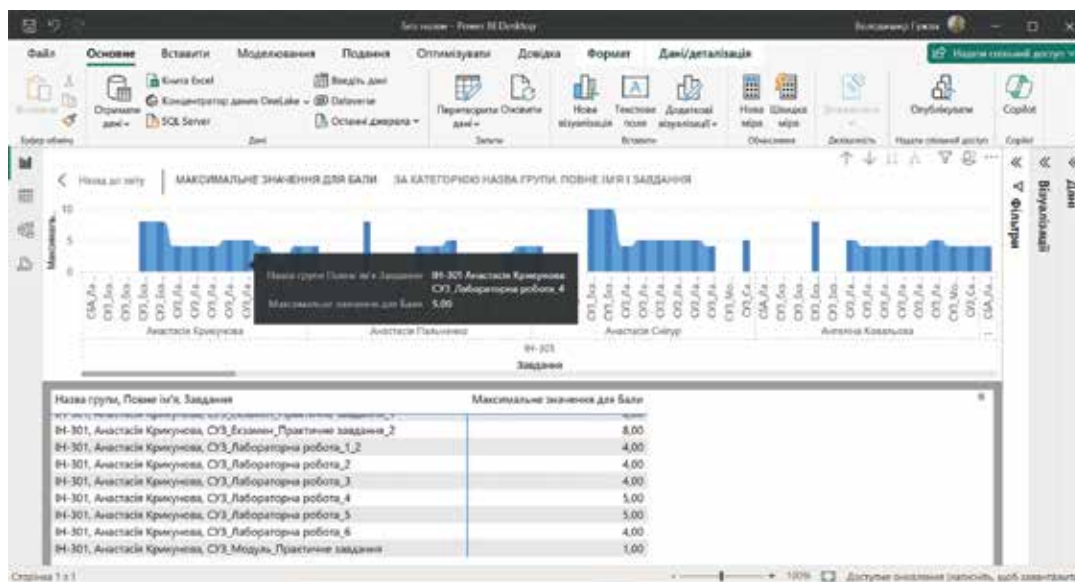
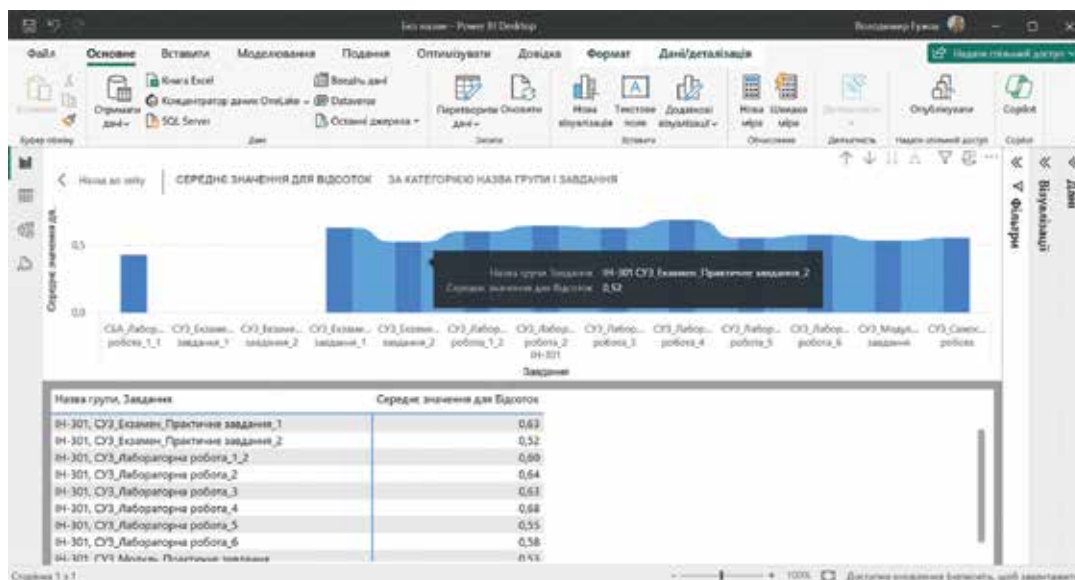


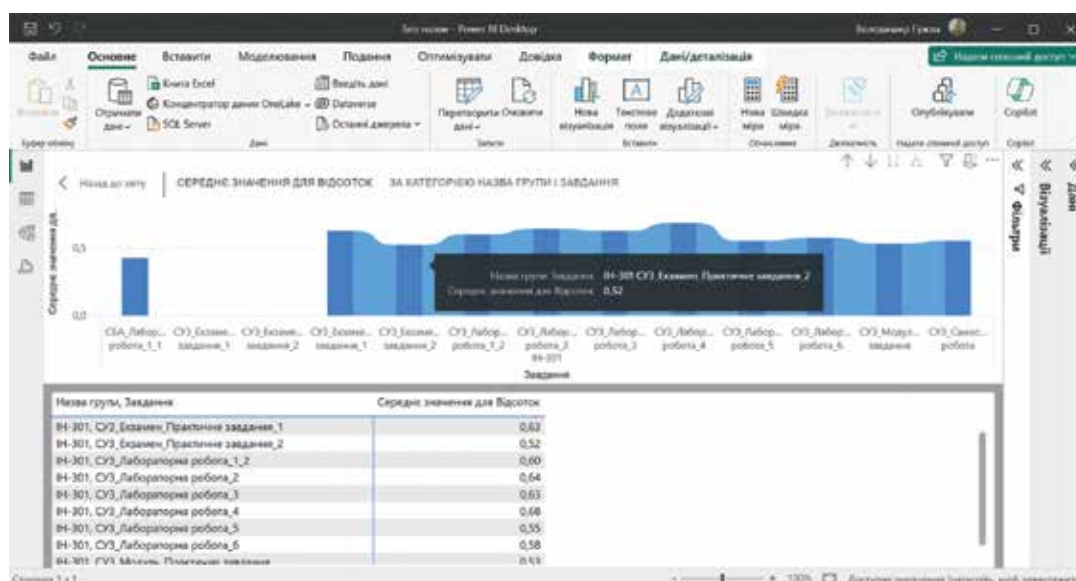
Рис. 2. Фрагмент електронного журналу успішності студентів (Grades) однієї із груп в модулі Teams



а)



б)



в)

Рис. 3. Результати проведеного EDM-аналізу за допомогою платформи Power BI

Джерело: розроблено автором

ють усі рівні управління та формують єдину базу знань. Водночас залишаються виклики: недостатня цифрова зрілість закладів, потреба у стандартизації методів і врахуванні етичних аспектів роботи з освітніми даними. Подальші дослідження доцільно спрямувати на мульт

тиалгоритмічний аналіз і тіснішу інтеграцію EDM із системами управління. Загалом впровадження EDM формує нову культуру управління освітою, де дані стають стратегічним ресурсом, а цифрова трансформація – ключем до сталого розвитку закладів.

References:

1. Díaz-Choque M., Chamorro-Atalaya O., Ortega-Galicio O. A., Arévalo-Tuesta J. A., Cáceres-Cayllahua E., Dávila-Laguna R. F., Aybar-Bellido I. E., Sigua-Jerónimo Y. B. (2023). Contributions of Data Mining to University Education, in the Context of the Covid-19 Pandemic: A Systematic Review of the Literature. *International Journal of Online and Biomedical Engineering (iJOE)*, no. 19(12), pp. 16–33. DOI: <https://doi.org/10.3991/ijoe.v19i12.40079>
2. Aldowah H., Al-Samarraie H. & Fauzy W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, no. 37, pp. 13–49. DOI: <https://doi.org/10.1016/J.TELE.2019.01.007>
3. R. Alarcón, J. Bravo, C. Valdivia and J. Aquino, "Analytical system based on Big Data for digital transformation in Higher Education," *2022 IEEE 2nd International Conference on Advanced Learning Technologies on Education & Research (ICALTER)*, Lima, Peru, 2022, pp. 1–4. DOI: <https://doi.org/10.1109/ICALTER57193.2022.9964613>
4. Chytas K., Tsolakidis A., Triperina E. & Skourlas C. (2022). Educational data mining in the academic setting: employing the data produced by blended learning to ameliorate the learning process. *Data Technol. Appl.*, no. 57, pp. 366–384. DOI: <https://doi.org/10.1108/DTA-06-2022-0252>
5. Antonio Fernández, Beatriz Gómez, Kleona Binjaku, Elinda Kajo Meçe. (2023) Digital transformation initiatives in higher education institutions: A multivocal literature review. *Education and Information Technologies*. No. 28. pp. 12351–12382. DOI: <https://doi.org/10.1007/s10639-022-11544-0>
6. Aulakh K., Roul R.K. & Kaushal M. (2023). E-learning enhancement through educational data mining with Covid-19 outbreak period in backdrop: A review. *International Journal of Educational Development*, no. 101, pp. 102814–102814. pp. 1–17.
7. Romero C. & Ventura S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, no. 10(3). DOI: <https://doi.org/10.1002/widm.1355>
8. García F., Valls C. y Lázaro J.L. (2022). Estrategias para la transformación digital de un centro educativo: una revisión sistemática. *RiiTE Revista interuniversitaria de investigación en Tecnología Educativa*, no. 13, pp. 157–172. DOI: <https://doi.org/10.6018/riite.533971>
9. Chrisnanto Y.H. & Abdullah G. (2021). The uses of educational data mining in academic performance analysis at higher education institutions (case study at UNJANI). *Matrix : Jurnal Manajemen Teknologi dan Informatika*. Pp. 26–35. DOI: <https://doi.org/10.31940/MATRIX.V11I1.2330>
10. Beck J.E., Woolf B.P. (2000). High-Level Student Modeling with Machine Learning. In: Gauthier G., Frasson C., Van-Lehn K. (eds) *Intelligent Tutoring Systems. ITS 2000. Lecture Notes in Computer Science*, vol 1839. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/3-540-45108-0_62
11. Tang T., McCalla G. (2004). Evaluating a Smart Recommender for an Evolving E-learning System: A Simulation-Based Study. In: Tawfik A.Y., Goodwin S.D. (eds) *Advances in Artificial Intelligence. Canadian AI 2004. Lecture Notes in Computer Science()*, vol 3060. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-24840-8_34
12. George Siemens and Ryan S. J. d. Baker. (2012). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12)*. Association for Computing Machinery, New York, NY, USA, pp. 252–254. DOI: <https://doi.org/10.1145/2330601.2330661>

13. Khine M. S. (2024). Educational Data Mining and Learning Analytics. In: Khine, M.S. (eds) Artificial Intelligence in Education. Springer, Singapore. DOI: https://doi.org/10.1007/978-981-97-9350-1_1
14. Xiao W., Ji P. & Hu J. (2022). A survey on educational data mining methods used for predicting students' performance. *Engineering Reports*, no. 4(5). DOI: <https://doi.org/10.1002/eng2.12482>
15. Barbeiro L., Gomes A., Correia F. B. & Bernardino J. (2023). A Review of Educational Data Mining Trends. *Procedia Computer Science*, no. 237, pp. 88–95. DOI: <https://doi.org/10.1016/j.procs.2024.05.083>
16. Papadogiannis I., Wallace M. & Karountzou G. (2024). Educational Data Mining: A Foundational Overview. *Encyclopedia*, no. 4(4), pp. 1644–1664. DOI: <https://doi.org/10.3390/encyclopedia4040108>
17. Lin Y., Chen H., Xia W., Lin F., Wang Z. & Liu Y. (2023). A Comprehensive Survey on Deep Learning Techniques in Educational Data Mining. *ArXiv*. DOI: <https://arxiv.org/abs/2309.04761>

Стаття надійшла: 29.08.2025

Стаття прийнята: 17.09.2025

Стаття опублікована: 26.09.2025