КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (27), 2025
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2025.27.709
УДК 004.056

Лахно Мирослав Валерійович
аспірант
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID ID: 0000-0001-6979-6076
valss725@gmail.com

СИСТЕМНИЙ АНАЛІЗ ЦИФРОВИХ СЛІДІВ У ІНФОРМАЦІЙНО-ОСВІТНІЙ СИСТЕМІ УНІВЕРСИТЕТУ

Анотація. В умовах стрімкого розвитку цифрових технологій і зростання кіберзагроз для інформаційних систем, які широко використовуються в багатьох галузях людської діяльності, актуальність засвоєння компетентностей з інформаційної безпеки (ІБ) стає все більш очевидною. Підготовка здобувачів вищої освіти (далі здобувачів ВО) закладів вищої освіта (далі ЗВО) у галузі ІБ не лише розширює їхні професійні компетентності, але й відіграє ключову роль у формуванні кваліфікованих спеціалістів, здатних ефективно вирішувати нові виклики у сфері кіберзагроз. Такі спеціалісти матимуть проактивну позицію й здатність до самоорганізації в умовах постійно змінюваних кіберзагроз, що особливо важливо в бізнес-процесах, які побудовані на цифрових технологіях. У цій роботі запропоновано алгоритм для системи підтримки ухвалення рішень (СПУР), спрямований на підвищення якості освіти й рівня захищеності цифрового освітнього середовища ЗВО (далі (ЦОС університету — ЦОСУ). Алгоритм заснований на аналізі цифрових слідів (ЦС) користувачів і може бути впроваджений в модель інтелектуального помічника для ЦОСУ. Суть запропонованого підходу полягає у використанні матричної факторизації ЦС користувачів, що дозволяє більш ефективно керувати й аналізувати інформацію про дії здобувачів ВО і науково-педагогічного/педагогічного персоналі у цифровому просторі. Однією з ключових переваг цього підходу є його здатність розв'язувати проблему розвитку компетентнісного профілю здобувачів ВО, особливо у сфері ІБ. Зазначений алгоритм сприяє більш глибокому розумінню й засвоєнню необхідних навичок, що зі свого боку значно підвищує ступінь захищеності ЦОСУ й комп’ютерних систем ЗВО. В умовах зростання кількості кіберзагроз і підвищення складності їхніх проявів, запропоновані рішення допомагають забезпечити надійний захист освітньої інфраструктури й сприяють підготовці спеціалістів, готових до викликів сучасного цифрового світу.

Ключові слова: цифрове середовище; заклад вищої освіти (ЗВО); цифрові сліди; матрична факторизація; машинне навчання (МН); інформаційна безпека; компетентності.

ВСТУП

Створення цифрового (у деяких вчених — цифрового освітнього) середовища в ЗВО (далі — ЦОСУ) — один із ключових елементів сучасної освітньої системи в будь-якій державі, яка прагне забезпечити якісну освіту. Цифровізація діяльності закладів освіти, зокрема ЗВО України в умовах агресії з боку росії породжує зростання обсягу даних, що вимагають подальшої обробки. Як було показано в роботах [1] – [3], з розвитком інформаційних технологій (ІТ) якісне ЦОСУ сприяє підвищенню якості навчання здобувачів ВО. Досвід використання освітнього контенту в електронних системах навчання, наприклад, таких, як Moodle, Blackboard, Canvas, Google Classroom, у період пандемії короновірусу COVID-19 показав, що здобувачі ВО можуть навчатися віддалено [4], [5].

Під час освітнього процесу здобувачі ВО й співробітники ЗВО залишають різні так звані цифрові сліди (далі — ЦС). На сьогодні ЦС займають значну частину хмари великих даних ЦОСУ, і важливим напрямком їхнього використання є вилучення інформації щодо активностей здобувачів ВО у період навчання. Наведемо перелік ЦС, які залишають здобувачі ВО в ході навчання і взаємодій з ЦОСУ [6] – [8]: електронна пошта; онлайн-платформи навчання; соціальні медіа; цифрові файли; інтернет-пошук та ін. Наведені ЦС можуть бути корисні для здобувачів ВО, науково-педагогічного/педагогічного персоналу, адміністраторів ЗВО при відстежуванні освітнього прогресу, оцінці робіт, комунікації й аналізі даних для підвищення якості освітнього процесу.

Отримувана з ЦС інформація може бути використана не лише в зовнішніх взаємодіях у ЦОСУ, але й у внутрішній діяльності ЗВО. Сучасні ЗВО все частіше використовують дані для підвищення задіяності науково-педагогічного/педагогічного персоналу у досягненні цілей і місії ЗВО, тобто для підвищення продуктивності як науково-педагогічного/педагогічного персоналу, так і діяльності ЗВО. Таким чином можна створювати нові джерела для своєї конкурентної переваги на ринку освітніх послуг.

Як було показано в роботах [9], [10], невід’ємною частиною передових ЗВО стали рекомендаційні системи, а також системи з елементами штучного інтелекту (ШІ), залучені до процесів управління ЗВО. Рекомендаційні системи, а також системи з елементами ШІ, наприклад, різні системи підтримки ухвалення рішень, базуються на ансамблі алгоритмів, програм і сервісів, призначених для формування релевантних рекомендацій для користувачів.

Зазначимо, що не менш важливим для будь-якого ЗВО є також завдання щодо забезпечення конфіденційності й захисту ЦС і персональних даних здобувачів ВО і науково-педагогічного/педагогічного персоналу. Останнє обумовлено тим, що ЦС користувачів ЦОСУ в багатьох випадках містять особисту інформацію. Отже, інформація, розміщена в ЦС, може бути об’єктом кібератак або зловживань. В умовах глобалізації освіти [11] ЗВО повинні максимально адаптувати свою політику інформаційної безпеки (ІБ) до захисту даних у ЦС, а також інформувати здобувачі ВО щодо стратегії ЗВО в завданнях захисту даних користувачів, політики його ІБ і конкретних практик захисту їхніх ЦС.

Інформаційна безпека ЦОСУ являє собою складну систему, яка передбачає захист наявного в ЗВО інформаційного простору. Така система запобігає пошкодженню або викрадення персональних даних учасників освітнього процесу, а також інформації, яка має фінансову, інтелектуальну цінність тощо. Забезпечення ефективного функціонування системи ІБ ЦОСУ передбачає витрати певних грошових ресурсів у межах обраної ЗВО стратегії захисту даних. При розробці подібної стратегії доцільно врахувати фактори зовнішнього і внутрішнього середовища, оскільки досягнення оптимального результату може бути забезпечене лише у випадку винайдення рівноваги між наявними можливостями й бажаними результатами. До таких результатів належить також інтегральна ціль — підвищення якості освіти на основі залучення потенціалу всіх форм організації освітнього процесу й розвитку інфраструктури ЗВО в умовах цифрової трансформації. за такої ситуації для менеджменту ЗВО й персоналу, відповідального за політику ІБ, може бути затребуваним контекстно-керований підхід до інтелектуальної підтримки ухвалення рішень щодо забезпечення ІБ ЦОСУ на основі аналізу ЦС користувачів. Усе вище зазначене й визначило наш інтерес до наукових досліджень у цьому напрямі.

Постановка проблеми. У сучасних умовах цифровізації вищої освіти особливої актуальності набуває проблема забезпечення ефективності та безпеки цифрового освітнього середовища ЗВО, оскільки зростаюча складність інформаційних систем та поширення кіберзагроз вимагають розроблення нових підходів до підтримки ухвалення рішень та моніторингу безпекової складової освітнього простору. Основним викликом є необхідність комплексного аналізу ЦС користувачів цифрового освітнього середовища з метою: підвищення якості освітнього процесу; посилення захисту інформаційних систем; розвитку компетентнісного профілю здобувачів ВО, особливо в питаннях інформаційної безпеки. Наявні системи підтримки ухвалення рішень мають обмежені можливості щодо ефективного використання інформації про ЦС користувачів та комплексного оцінювання стану захищеності освітнього середовища, а це, у свою чергу, обумовлює необхідність розроблення нових алгоритмічних підходів, здатних інтегрувати методи інтелектуального аналізу даних та забезпечити підвищення рівня ІБ ЗВО.

Аналіз останніх досліджень і публікацій. В роботах [12], [13] автори розглядають питання аналізу ЦС здобувачів ВО в завданнях застосування відповідних інтелектуальних систем з елементами машинного навчання (далі МН) в ЗВО. Зокрема, розглянуто аспекти цієї задачі, котрі тісно пов’язані з підвищенням ефективності освітнього процесу й покращенням результатів навчання.

У роботі [14] автори розглядають проблеми етапу збору даних на основі ЦС, починаючи з академічних показників с здобувачів ВО і закінчуючи поведінковими моделями здобувачів. На думку дослідників, подібна інформація може слугувати основою при створенні інформаційної бази, яка буде використовуватися відповідними системами.

У працях [15], [16] науковці торкаються проблеми аналізу й обробки даних на основі ЦС здобувачів освіти. Автори робіт [17], [18] узагальнили моделі, використовувані в ході аналізу й обробки ЦС здобувачів освіти, що дозволило за допомогою ІТ виявляти певні закономірності, залежності й тенденції процесу навчання. У результаті подібний аналіз, на думку дослідників, дозволить оптимізувати освітній процес і надавати персоналізовані рекомендації для кожного здобувача освіти.

Однак зазначимо, що переважна більшість проаналізованих робіт зовсім не торкається аспектів, пов’язаних з аналізом ступеня захищеності ЦОСУ на основі вивчення ЦС користувачів. Виходячи з вище зазначеного, актуальною метою дослідження є створення інтелектуального помічника (ІП) з елементами МН, який буде супроводжувати здобувача освіти в період навчання, персоніфікуючи освітню траєкторію на основі даних, складених за ЦС щодо особистого профілю й реакцій на освітній контент, зокрема щодо ІБ ЦОСУ.

Мета дослідження — розробка та тестування алгоритму для системи підтримки ухвалення рішень (СПУР) у цифровому освітньому середовищі ЗВО, який дозволяє підвищити якість освіти та рівень захищеності цього середовища.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

В основу розробки ІП (див. рис. 1) покладено принципи системної інженерії. Концептуально ІП базується на модульній архітектурі. Основне завдання діагностичного модуля — формування цифрового профілю користувача (ЦПК), перш за все здобувача ВО, на основі аналізу ЦС в ЦОСУ. Такий ЦПК дозволить максимально враховувати всі його особливості для формування ефективної освітньої траєкторії, яка сприятиме підвищенню якості освіти.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (27), 2025
ISSN 2663 - 4023

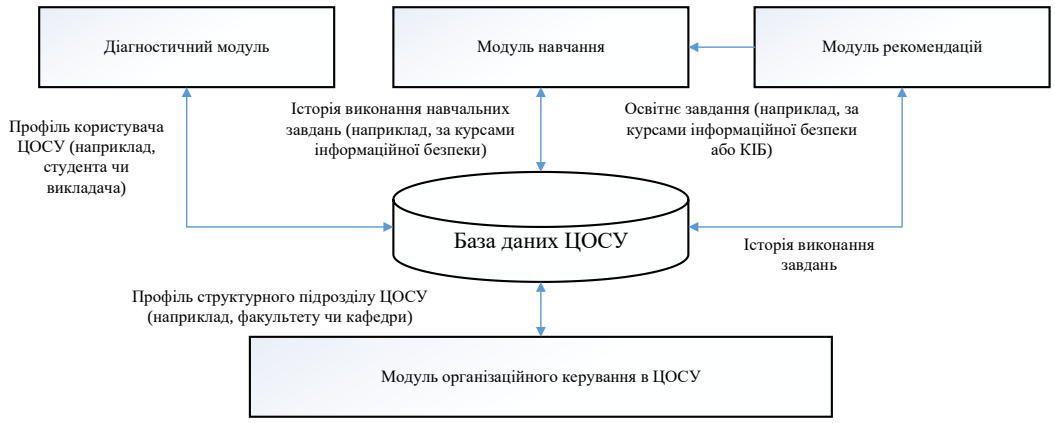

Рис. 1. Концептуальна модель інтелектуального помічника для ЦОСУ

За умов, коли бізнес-процеси багатьох компаній базуються на широкому використанні ІТ, роботодавці звертають пильну увагу не лише на традиційні компетентності, закладені в освітніх програмах ЗВО за відповідними спеціальностями, але й приділяють пильну увагу загальній цифровій грамотності випускників ЗВО. А базовий набір цифрових компетентностей в умовах зростання кількості кібернетичних загроз на сьогодні розширений за рахунок компетентностей, які стосуються знань і навичок з дотримання інформаційної безпеки як особистих даних, так і корпоративних.

За умов, коли бізнес-процеси багатьох компаній базуються на широкому використанні ІТ, роботодавці звертають пильну увагу не лише на традиційні компетентності, закладені в освітніх програмах ЗВО за відповідними спеціальностями, але й приділяють пильну увагу загальній цифровій грамотності випускників ЗВО. А базовий набір цифрових компетентностей в умовах зростання кількості кібернетичних загроз на сьогодні розширений за рахунок компетентностей, які стосуються знань і навичок з дотримання інформаційної безпеки як особистих даних, так і корпоративних.

У створюваному ЦПК ЦОСУ повинні враховуватись не лише персональні дані користувача, але й дані, що відображають рівні розвитку його компетентностей, зокрема soft і hard skills, які включають компетентності з ІБ.

На модуль організаційного керування покладено завдання щодо створення організаційної структури ЦОСУ.

Основний модуль ІП — це модуль організації освітнього процесу. Для кожного завдання, що вирішується в ході навчання в ЦОСУ формується ЦС, який дозволить викладачам (науково-педагогічному/педагогічному персоналу) і адміністрації ЗВО скласти уявлення про те, наскільки для здобувача освіти запропоноване завдання було цікавим, складним, релевантним, а також наскільки ефективним було його виконання. Користувачі ЦОСУ мають можливість оцінити, наскільки сподобалось їм завдання. Отже, формується ЦС у вигляді наборів даних, котрі характеризують виконання користувачами освітніх завдань, зокрема завдань, пов’язаних з питаннями ІБ ЦОСУ. Такі набори можуть у подальшому використовуватись при формуванні релевантних рекомендацій для менеджменту ЗВО.

Модуль рекомендацій ІП слугує для формування індивідуальної освітньої траєкторії здобувача освіти, спираючись на його пріоритети й цінності, стиль і темп навчання. Відповідно на підставі цих даних система може підібрати максимально релевантний контент.

Процедуру підбору релевантного освітнього контенту можна вирішити безпосередньо за допомогою рекомендаційної системи. У якості алгоритму роботи такої

системи була задіяна технологія МН, яка побудована на матричній факторизації. Далі в роботі детальніше зупинимось на цьому.

При аналізі ЦС в ЦОСУ адміністраторам, а також спеціалістам з ІБ найчастіше доводиться стикатися з аналізом лог-файлів. Причому лог-файли й ЦС з'єднані в тлумаченні аналізу цифрових доказів дій користувачів. Наведемо лише деякі приклади взаємозв'язку лог-файлів і ЦС користувачів у ЦОСУ: зберігання інформації в ЦОСУ; відновлення подій; аналіз даних тощо. Отже, лог-файли й ЦС взаємопов'язані, оскільки лог-файли містять інформацію про події й дії, котрі можуть бути проаналізовані для виявлення, ідентифікації й інтерпретації ЦС для аналізу цифрових доказів активностей користувачів ЦОСУ як в завданнях підвищення якості освітнього процесу, так і для ІБ ЦОСУ. По суті лог-файлы — це незалежні характеристики роботи користувача в мережі ЗВО. Вони утримують інформацію щодо входів до системи, використання ресурсів, помилок, мережевої активності й інших подій у ЦОСУ.

Однак, щоб повністю зрозуміти контекст роботи користувача, лог-файли зазвичай потребують аналізу й інтерпретації з використанням інших інструментів. Контекстно залежні характеристики роботи користувача в ЦОСУ можуть включати інформацію про час, місцезнаходження, використані додатки й інші фактори, пов'язані з конкретною ситуацією або завданням користувача в ЦОСУ. Ці характеристики можуть бути отримані з лог-файлів. Однак, зазвичай потрібен додатковий аналіз. Такий аналіз може бути реалізований за допомогою спеціалізованого програмного забезпечення (ПЗ), зокрема Splunk, ELK Stack (стек програмних продуктів, що містить Elasticsearch, Logstash і Kibana) тощо [19], [20]. Підвищити ступінь захищеності ЦОСУ можна зокрема шляхом використання систем підтримки ухвалення рішень (СПУР), ШІ й методів МН.

У розрізі цілей дослідження необхідно розробити алгоритм для СПУР, який би сприяв підвищенню якості освіти й ступеня захищеності ЦОСУ на підставі аналізу ЦС. У пропонованому нижче в роботі підході ШІ виражений за допомогою методів МН з використанням матричної факторизації.

На нашу думку, в ЦОСУ є значна кількість користувачів: $U = \{u_1,...,u_n\}$ і значна кількість як освітніх завдань $E = \{e_1,...,e_m\}$, так і завдань, пов'язаних з ІБ ЦОСУ — $S = \{s_1,...,s_k\}$. Тоді адміністратори ЦОСУ мають доступ до матриць, які містять, наприклад, оцінки, виставлені користувачами за їхньою пріоритетністю для освітніх завдань — $ME_{n\times m}$, а також такі., що характеризують користувача з точки зору дотримання правил ІБ при роботі в ЦОСУ — $MS_{n\times k}$.

У матриці $ME_{n\times m}$ на місці me_{ij} $(i \in 1,...,n;\ j \in 1,...,m)$ ставиться певне число, якщо користувач ЦОСУ (u_i) оцінює завдання (e_j), виходячи зі своїх пріоритетів, і залишається порожнім в іншому випадку. Дані беруться на основі ЦС користувачів із Moodle, Blackboard, Canvas, Google Classroom тощо.

Для заповнення другої матриці за критеріями безпечної поведінки в ЦОСУ виділимо декілька типів користувачів за рівнем їхніх компетентностей у питаннях ІБ.

Обізнані користувачі. До цієї групи відносимо користувачів, котрі добре обізнані про ризики ІБ в мережі ЗВО або ЦОСУ в цілому. Такі користувачі уживають активних заходів для забезпечення ІБ своїх даних і акаунтів у ЦОСУ, пунктуально виконують рекомендації щодо створення складних паролів, регулярно оновлюють ПЗ, не відкривають підозрілі посилання чи вкладки в електронних листах і використовують надійне антивірусне ПЗ.

Недбалі користувачі. До цієї групи відносимо користувачів, котрі не звертають належної уваги на заходи ІБ. Відповідно вони досить вразливі до атак, як зовнішніх, так і внутрішніх. Для користувачів цієї групи характерне використання слабких паролів, їхній повтор для різних акаунтів. Ці користувачі, як правило, ігнорують підозрілу активність у мережі й не дотримуються заходів щодо захисту своїх даних в ЦОСУ.

Необізнані користувачі. До цієї групи відносимо користувачів, котрі не володіють достатнім рівнем знань про заходи з ІБ під час роботи в мережі. Вони можуть і не знати про ризики, пов'язані з відкриттям підозрілих посилань, не знайомі з правилами використання загальнодоступних Wi-Fi мереж в ЦОСУ. Такі користувачі без застережень встановлюють ненадійне ПЗ, часто передають конфіденційні дані через незахищені канали зв'язку.

Байдужі користувачі. До цієї групи відносимо користувачів, котрі не виявляють ніякого інтересу до питань ІБ в мережі ЗВО й не дотримуються елементарних заходів з ІБ.

Безвідповідальні користувачі. До цієї групи відносимо користувачів ЦОСУ, котрі порушують правила й політики ІБ в мережі ЗВО. Вони можуть намагатись отримати несанкціонований доступ до систем ЦОСУ, розповсюджувати шкідливі програми, порушувати конфіденційність даних чи вести недобросовісну активність у ЦОСУ.

Наведена вище категоризація типів користувачів досить умовна. Як зазначено в праці [21], відсутні чіткі межі між згаданими категоріями користувачів. У міру набуття знань, наприклад, за рахунок відповідних компонентів в освітніх програмах ЗВО, користувачі можуть переходити від одного типу до іншого, усвідомлюючи важливість ІБ в мережі й застосовуючи певні заходи для захисту своїх даних і акаунтів. Тоді в матриці $MS_{n\times k}$ на місці $ms_{ij}\ (i \in 1,...,n;\ j \in 1,...,k)$ ставиться певне число, якщо користувач ЦОСУ (u_i) віднесений до певної групи (s_j), виходячи зі стилю своєї поведінки в ЦОСУ в розрізі дотримання правил ІБ. Фактично в цій матриці відображені дані, що стосуються компетентностей з ІБ здобувачів освіти й співробітників. В цифрованому вигляді подібна матриця може містити, наприклад, стиль поведінки в питаннях ІБ, здобувача ВО або співробітника. Такі дані отримані на підставі аналізу ЦС, наприклад, з використанням методики, наведеної в працях [22], [23]. Місце залишається порожнім в іншому випадку, тобто, коли оцінка стилю за ІБ не виконана.

Потрібно знайти вектори $\left(\hat{me}_i\right)$, $\left(\hat{ms}_i\right)$, які містять такі дані:

1) в завданнях формування індивідуальної освітньої траєкторії вже відомі оцінки користувача (u_i), тобто $\left(\hat{me}_i\right)$. Також передбачувані оцінки — $\left(\hat{me}_{ij}\right)$;
2) в завданнях формування навичок безпечної роботи в ЦОСУ вже відомі навички, наприклад, на основі ЦС або результатів тестування, тобто $\left(\hat{ms}_i\right)$.

   Також передбачувані оцінки після отримання відповідних компетентностей з ІБ — $\left(\hat{ms}_{ij}\right)$.

Оскільки одним із завдань дослідження була розробка алгоритму для СПУР як методу МН, було використано матричну факторизацію (МФ). Під МФ мають на увазі декомпозицію вихідної матриці в добуток двох матриць малого рангу [24], [25]. Відповідно взаємодія користувачів з об'єктом буде моделюватися як скалярний добуток векторів уявлення користувачів і об'єктів у факторному просторі, що стосується, наприклад, компетентностей здобувачів ВО у питаннях ІБ. Зазначимо, що

факторизаційні моделі добре зарекомендували себе в роботі з сильно розрядженими матрицями [24], [25]. Це пов'язано з тим, що МФ дозволяє отримувати приховані залежності на основі аналізу ЦС користувачів у ЦОСУ й робити прогнози на основі великих обсягів інформації, циркулюючої в будь-якому ЗВО.

Для МН СПУР МФ може застосовуватись, наприклад, для завдань вироблення рекомендацій, пов'язаних як з підвищенням якості освітнього процесу в цілому, так і окремих компетентностей здобувачів ВО і співробітників, наприклад, з ІБ.

Оскільки робота з матрицями аналогічна, у межах нашого дослідження ми розглядаємо лише алгоритм роботи з матрицею оцінок здобувачів ВО у ЦОСУ (див. табл. 1 і рис. 2).

1) При збільшенні обсягів отриманих даних, отже, зменшенні розрідженості даних, може знадобитись збільшення числа ітерацій $(iter)$. Контроль числа ітерацій $(iter)$ може бути автоматизований. На сьогодні здійснюється розробка відповідного ПЗ, за допомогою якого можна буде після закінчення факторизації й порівняння точності передбачень з попередніми результатами виробити за допомогою СПУР з покращення якості освітнього процесу, зокрема компетентностей з ІБ. Якщо точність передбачень зменшиться, слід збільшити число ітерацій. В іншому випадку число ітерацій не зміниться (див. рис. 2). Як альтернативу можна задати допустиму точність факторизації. Якщо ця точність досягнута, то робота алгоритму, показаного на рис. 2, зупиняється.

Таблиця 1

Кроки алгоритму пошуку рішення мінімізації помилки при МН на основі матричної факторизації даних аналізу ЦС користувачів ЦОСУ

Номер кроку	Опис дії	Математична інтерпретація й розшифрування параметрів
1	Знаходимо матрицю B	
2	Визначимо помилку δ	$\delta = \left\| me - \hat{me} \right\|,\ j \in 1,...,n$
3	Знаходимо нові значення A_{ir}	$A_{ir} = A_{ir} - \nu\left(\delta B_{rj}^{E} + \lambda A_{ir}\right)$, де $r \in 1,...,k,\ \lambda -$ регуляційний параметр; $\nu -$ швидкість навчання
4	Знаходимо матрицю A	
5	Визначимо помилку δ	$\delta = \left\| me - \hat{me} \right\|,\ j \in 1,...,m$
6	Знаходимо нові значення B_{rj}	$B_{rj} = B_{rj} - \nu\left(\delta A_{ir}^{E} + \lambda B_{rj}\right)$

Представимо матрицю оцінок $ME_{n\times m}$ у вигляді добутку двох матриць:

1) матриці $A_{n\times w}$, котра містить числовий опис прихованих (латентних) характеристик користувачів (наприклад, поведінкові патерни: регулярність активності в ЦОСУ, частота й час входу до ЦОСУ, типові інтервали активності; споживання контенту; рівень доступу, частота помилкових входів до системи, спроби доступу до заборонених ресурсів, а також явних характеристик користувачів (курс, вік, середні оцінки тощо);

2) матриці $B_{w\times u}$, котра характеризує освітні завдання, наприклад, пріоритетність курсів з ІТ й/або ІБ для формування індивідуальної освітньої траєкторії.

Заповнюємо випадковими величинами на основі закону рівномірного розподілу на интервалі $\left\lfloor 0;\ \sqrt{\max\left\{me_{ij}\right\}/k}\right\rfloor$ латентні характеристики відповідно для матриць A і B.

Потім вирішуємо задачу мінімізації, скориставшись залежністю (1):

$$\arg\min \left\| ME - \hat{ME} \right\| + \alpha \|B\| + \beta \|A\|, \quad (1)$$

де $\hat{ME}$ — матриця, отримана в результаті апроксимації з матриць A і B; α, β — параметри алгоритму.

На кожному кроці ітераційного алгоритму мінімізація помилок буде включати кроки, наведені в табл. 1 і на рис 2.

Експериментальні дослідження проводились на базі Національного університету біоресурсів и природокористування України. Щоб оцінити якість передбачень ІП (див. рис. 1) більшість оцінок здобувачів за освітніми компонентами, пов'язаними з ІБ (освітні компоненти для різних спеціальностей мають різні назви, наприклад: «Основи захисту інформації», «Основи інформаційної безпеки», «Технології захисту інформації» тощо) $ME_{n \times m}$ було розділене на вибірки. Це відповідно вибірка для навчання — ME_{training} і вибірка для тестування — ME_{test}.

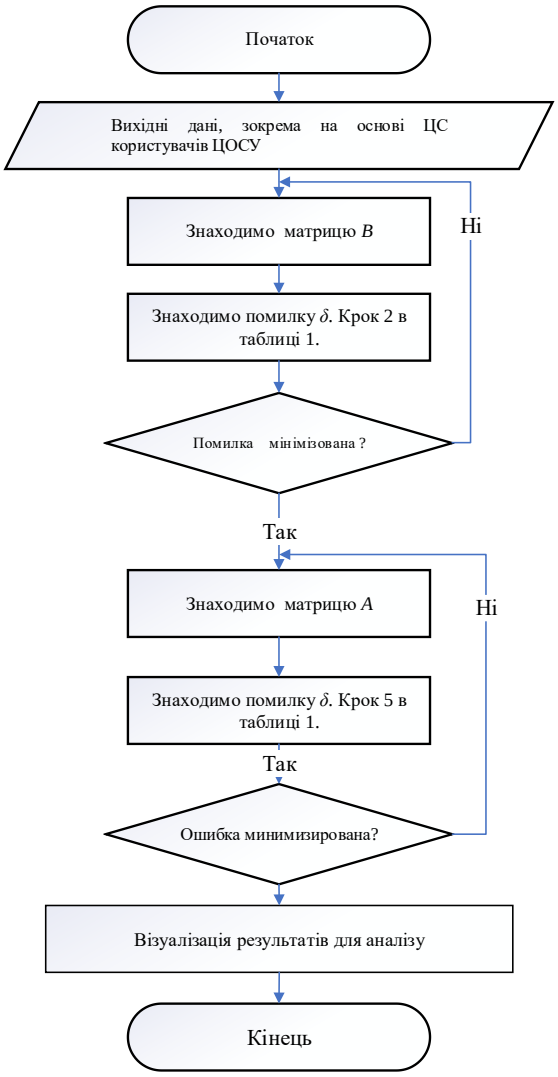

Рис. 2. Блок-схема алгоритму матричної факторизації на основі аналізу даних ЦС користувачів у ЦОСУ

Тестування моделі, описаної раніше в цій роботі, було проведене за кількісних характеристик, представлених у табл. 2.

Оцінювання роботи моделі МН ІП для ЦОСУ було проведене з допомогою трьох типів апроксимацій для: ефективності рішень (Ef); складності завдання з компонента ІБ $(ComT)$; уподобань у стилі ставлення до ІБ здобувачів освіти $(\Pr ef)$. Ми розглядали такі випадки:

1) використання тільки прихованих (латентних) характеристик здобувачів ВО (наприклад, поведінкові патерни на основі ЦС в ЦОСУ: регулярність активності в ЦОСУ, частота й час входу в ЦОСУ, типові інтервали активності; споживання контенту з ІБ; рівень доступу, частота помилкових входів до системи, спроби доступу до заборонених ресурсів тощо). Позначення на графіку — Latent;
2) використання явних і апріорі зафіксованих в ЦОСУ факторів: профіль здобувача ВО, курс, наявність hard skills з ІБ, наприклад, завантажені сертифікати з курсів, пов'язаних з ІБ, отримані на таких платформах, як Coursera, EdX тощо. Позначення на графіку — Explicit;
3) використання фіксованих факторів для завдань з ІБ. Ці фактори поряд з прихованими дозволяють оцінювати середні значення (Ef), $(ComT)$, $(\Pr ef)$. Позначення на графіку — Fixed;
4) використання фіксованих факторів і типів користувачів за рівнем їхніх компетентностей з питань ІБ. Позначення на графіку — Explicit+Fixed.

Як зазначено вище, робота з другою матрицею, що стосується категоризації типів користувачів з питань ІБ, аналогічна.

Таблиця 2

Кількісні характеристики експериментальних досліджень ІП для ЦОСУ

№	Назва характеристики	Значення
1	Кількість користувачів, котрі брали участь у тестуванні, людей	450
2	Кількість завдань, для яких відстежувались результати роботи (по 20 завдань, зокрема тестів у межах компонентів з ІБ)	9000
3	Кількість завдань, які були виконані користувачами	8996
4	Навчальна вибірка, %	75
5	Тестова вибірка, %	25
6	Ранг Матриць $A_{n\times w}$ и $B_{w\times u}$	40
7	Кроки факторизації матриць (процес розкладу матриці на добуток інших матриць)	25
8	Нормалізовані дані для діапазону значень вихідної матриці $ME_{n\times m}$	$0\div10$

Алгоритм пошуку рішення мінімізації помилки при МН на основі матричної факторизації даних аналізу ЦС користувачів ЦОСУ були програмно реалізований алгоритмічною мовою Python. У табл. 3 й на рис. 3, 4 відображено результати тестування алгоритму.

Як відомо, при МН може виникнути проблема холодного старту, тобто ситуація, коли нові дані або нові об'єкти з'являються в системі, а модель не може з ними коректна працювати через недостатність інформації [26], [27]. Для вирішення проблеми холодного старту можуть застосовуватись різні техніки. Наприклад, можна застосовувати метадані, контентні характеристики чи генеративні моделі для створення початкових передбачень. Також важливо постійно оновлювати й донавчати модель, використовуючи нові дані,

ISSN 2663 - 4023

щоб мінімізувати вплив проблеми холодного старту. Відповідно до праць [27], [28] ми вирішували цю проблему, застосувавши косинусну схожість [29], [30]:

$$\cos\varphi = \frac{\left(\overline{A_i}, \overline{B_j}\right)}{\left\|\overline{A_i}\right\| \times \left\|\overline{B_j}\right\|}.$$

Для цього знаходили косинус між новим користувачем, зареєстрованим в ЦОСУ, і кожним існуючим користувачем. Потім виконували ранжування за спаданням. У підсумку обирали користувача, зареєстрованого в ЦОСУ, якому відповідає максимальне значення косинусу [30]. Як результат, на підставі знайденого через косинус-схожість користувачів для нових користувачів будуть сформовані рекомендації, зокрема в питаннях, які стосуються підвищення їхніх компетентностей з ІБ.

Таблиця 3

Дані, отримані в ході тестування алгоритму матричної факторизації на основі аналізу даних ЦС користувачів у ЦОСУ

Швидкість навчання (ν) (Час, необхідний для завершення процесу навчання моделі на основі матричної факторизації)	**Число ітерацій**		
	20	**40**	**60**
	Час навчання (τ), с; Середня абсолютна помилка факторизації матриць (δ)		
$\nu = 0{,}005$	$\tau = 6;\ \delta = 0{,}2$	$\tau = 11;\ \delta = 0{,}11$	$\tau = 16;\ \delta = 0{,}075$
$\nu = 0{,}015$	$\tau = 6;\ \delta = 0{,}12$	$\tau = 10;\ \delta = 0{,}06$	$\tau = 15;\ \delta = 0{,}03$
$\nu = 0{,}03$	$\tau = 5;\ \delta = 0{,}05$	$\tau = 11;\ \delta = 0{,}02$	$\tau = 15;\ \delta = 0{,}014$
$\nu = 0{,}04$	$\tau = 5;\ \delta = 0{,}048$	$\tau = 11;\ \delta = 0{,}022$	$\tau = 15;\ \delta = 0{,}014$

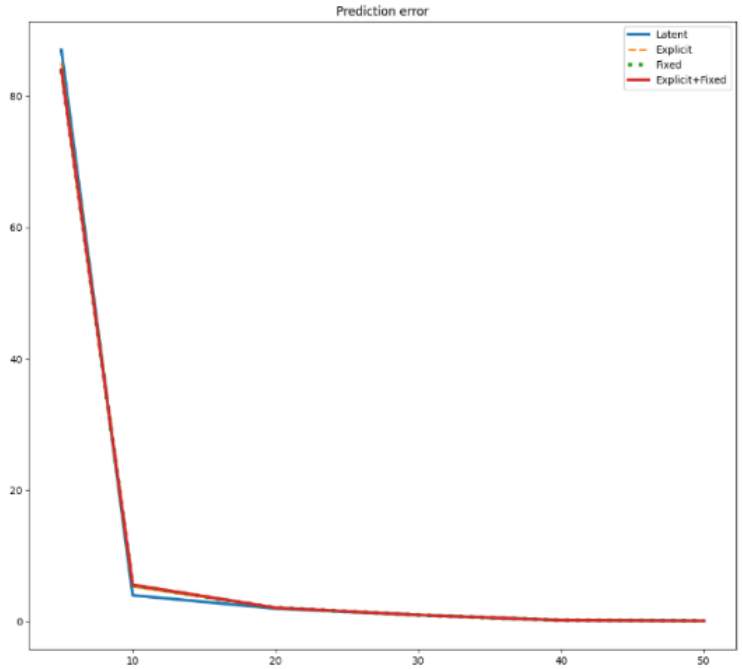

Рис. 3. Помилка передбачення уподобань здобувачів ВО при обранні дисциплін з ІБ за допомогою факторизації на відсоткових інтервалах

ISSN 2663 - 4023

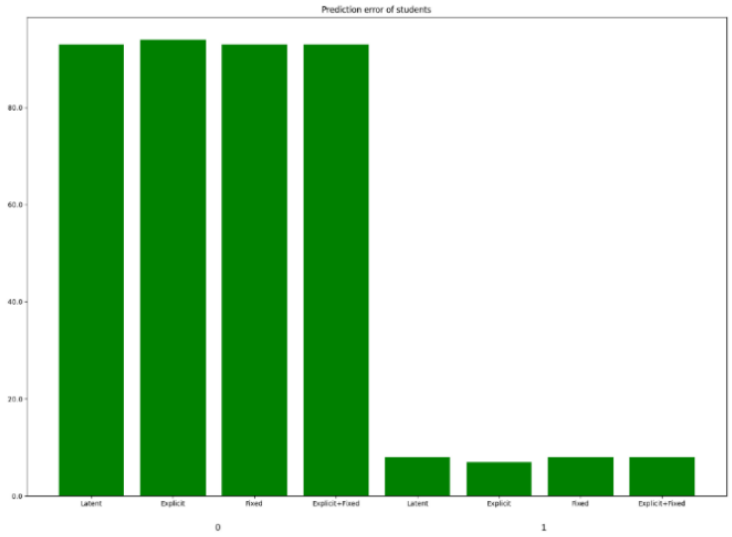

Рис. 4. Помилка передбачення уподобань здобувачів ВО при обранні дисциплін з ІБ за допомогою факторизації на дискретних величинах [0; 1]

Як видно з табл. 3, збільшення швидкості навчання ІП зі значення $v = 0{,}005$ до $v = 0{,}025$ дозволить зменшити похиби навчання. Відповідно зростає точність передбачення успішності навчання за освітніми компонентами, пов'язаними з ІБ. При $v = 0{,}04$ результати можна порівняти з результатами при $v = 0{,}03$. На заданих вибірках для числа ітерацій $iter = 20$ при $v = 0{,}03$ була зафіксована висока точність факторизації.

Збільшення обсягів даних щодо ЦС користувачів ЦОСУ, відповідно при руйнуванні розрідженості матриць, вимагає збільшення числа ітерацій при роботі алгоритму. Моніторинг числа ітерацій був реалізований також за допомогою Python шляхом виконання тестування після завершення процедури факторизації й порівняння точності передбачення оцінок здобувачів ВО з попередніми тестуваннями. Якщо точність передбачень падала, то число ітерацій збільшувалось.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У ході дослідження були отримані такі результати:

– запропоновано алгоритм для системи підтримки ухвалення рішень (СПУР) як елемента цифрового освітнього середовища ЗВО (ЦОСУ), який сприятиме підвищенню якості освіти й ступеня захищеності ЦОСУ. Важливим компонентом даних для ухвалення рішень є цифрові сліди (ЦС) користувачів ЦОСУ. Алгоритм застосовується як частина обчислювального ядра в моделі інтелектуального помічника для ЦОСУ і заснований на матричній факторизації ЦС користувачів. На відміну від відомих рішень, запропонований підхід дозволяє нівелювати проблему розвитку компетентнісного профілю здобувачів ВО, перш за все в питаннях придбання навичок з ІБ. У цілому це сприятиме збільшенню ступеня захищеності ЦОСУ й комп'ютерних систем ЗВО;

– запропонований алгоритм був реалізований мовою Python і були проведені експериментальні дослідження щодо тестування алгоритму матричної факторизації на основі аналізу даних ЦС користувачів ЦОСУ.

Перспективи подальших досліджень включають вдосконалення алгоритму для інтеграції з іншими компонентами ЦОСУ, а також розробку методів оптимізації його роботи на основі нових даних та технологій. Крім того, важливим напрямом є розширення застосування алгоритму для інших аспектів освітнього процесу, таких як оцінка ефективності навчальних програм і підвищення рівня індивідуалізації освітніх траєкторій здобувачів ВО.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Shevchuk, B. (2023). Virtualne osvitnie seredovyshche zakladu vyshchoi osvity: realii ta perspektyvy. *Nauka i tekhnika sohodni, 14(28),* 493–504.
2. Buinytska, O. P., Varchenko-Trotsenko, L. O., & Hrytseliak, B. I. (2020). Tsyfrovizatsiia zakladu vyshchoi osvity. *Osvitolohichnyi dyskurs: elektronne naukove fakhove vydannia, (28),* 64–79.
3. Areshonkov, V. Yu. (2020). Tsyfrovizatsiia vyshchoi osvity: vyklyky ta vidpovidi. *Visnyk Natsionalnoi akademii pedahohichnykh nauk Ukrainy, 2(2),* 1–6.
4. Pozo, J. I., Pérez Echeverría, M. P., Cabellos, B., & Sánchez, D. L. (2021). Teaching and learning in times of COVID-19: Uses of digital technologies during school lockdowns. *Frontiers in Psychology, 12, 656776.*
5. Mustapha, I., Van, N. T., Shahverdi, M., Qureshi, M. I., & Khan, N. (2021). Effectiveness of digital technology in education during COVID-19 pandemic. *A bibliometric analysis,* 136–154.
6. Vaccari, A., Calabrese, F., Liu, B., & Ratti, C. (2009, November). Towards the SocioScope: an information system for the study of social dynamics through digital traces. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 52–61.
7. Osborne, N., & Connelly, L. (2015). Managing your digital footprint: Possible implications for teaching and learning. In *Proceedings of the 2nd European Conference on Social Media ECSM*, 354–361.
8. Morze, N., Kuzminska, O., & Mazorchuk, M. (2019). Attitude to the digital learning environment in Ukrainian universities. In *ICT in Education, Research, and Industrial Applications, vol. 2393*, 53–67.
9. Owoc, M. L., Sawicka, A., & Weichbroth, P. (2019, August). Artificial intelligence technologies in education: benefits, challenges and strategies of implementation. In *IFIP International Workshop on Artificial Intelligence for Knowledge Management*, 37–58.
10. Ahmad, S. F., Alam, M. M., Rahmat, M. K., Mubarik, M. S., & Hyder, S. I. (2022). Academic and administrative role of artificial intelligence in education. *Sustainability, 14(3), 1101.*
11. Goodwin, A. L. (2020). Globalization, global mindsets and teacher education. *Action in Teacher Education, 42(1),* 6–18.
12. Ren, Z., Xin, Y., Ge, J., Zhao, Z., Liu, D., Ho, R. C., & Ho, C. S. (2021). Psychological impact of COVID-19 on college students after school reopening: a cross-sectional study based on machine learning. *Frontiers in Psychology, 12, 641806.*
13. Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments, 9(1), 11,* 2–19.
14. Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education, 44(1),* 130–160.
15. Van den Beemt, A., Groothuijsen, S., Ozkan, L., & Hendrix, W. (2023). Remote labs in higher engineering education: engaging students with active learning pedagogy. *Journal of Computing in Higher Education, 35(2),* 320–340.
16. Hasibuan, M., Reynolds, M., Male, S., & Hassan, G. M. (2023). Role of Theory in Analysing the dynamic of Self-regulated Learning process based on students' event logs data: A scoping review. *ASCILITE Publications*, 126–136.
17. Nevin, A. D., Chen, Y., Yang, S., & Quan-Haase, A. (2022). Key considerations in the interpretation of digital trace data. *The SAGE handbook of social media research methods*, 54–66.
18. Hüllmann, J. A. (2019). The construction of meaning through digital traces. *Proceedings of the Pre-ICIS*, 1–5.

19. Aarthi, M. (2021). Using Users Profiling to Identifying an Attacks. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12 (7)*, 795–802.
20. Subramanian, K., & Subramanian, K. (2020). Introducing the Splunk Platform. *Practical Splunk Search Processing Language: A Guide for Mastering SPL Commands for Maximum Efficiency and Outcome*, 1–38.
21. Shu, K., Zhou, X., Wang, S., Zafarani, R., & Liu, H. (2019, August). The role of user profiles for fake news detection. In *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining*, 436–439.
22. Sen, I., Flöck, F., Weller, K., Weiß, B., & Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. *Public Opinion Quarterly, 85(S1)*, 399–422.
23. Coulter, R., Han, Q. L., Pan, L., Zhang, J., & Xiang, Y. (2019). Data-driven cyber security in perspective – Intelligent traffic analysis. *IEEE transactions on cybernetics, 50(7)*, 3081–3093.
24. Lei, C., Dai, H., Yu, Z., & Li, R. (2020). A service recommendation algorithm with the transfer learning based matrix factorization to improve cloud security. *Information Sciences, 513*, 98–111.
25. Zheng, X., Guan, M., Jia, X., Guo, L., & Luo, Y. (2022). A matrix factorization recommendation system-based local differential privacy for protecting users' sensitive data. *IEEE Transactions on Computational Social Systems, 10(3)*, 1189–1198.
26. Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms: an overview. *Journal of physics: conference series, 1142*.
27. Fedoruk, P. I., Pikuliak, M. V., & Dutchak, M. S. (2010). Intelektualnyў mekhanizm pobudovy indyvidualnoï navchalnoï traiektoriï v adaptyvnykh systemakh dystantsiўnoho navchannia. *Shtuchnyi intelekt, 3*, 668–678.
28. Joy, J., & Renumol, V. G. (2020, December). Comparison of generic similarity measures in E-learning content recommender system in cold-start condition. In *2020 IEEE Bombay section signature conference (IBSSC)*, 175–179.
29. Xia, P., Zhang, L., & Li, F. (2015). Learning similarity with cosine similarity ensemble. *Information sciences, 307*, 39–52.
30. Mana, S. C., & Sasipraba, T. (2021, March). Research on cosine similarity and pearson correlation based recommendation models. *In Journal of Physics*: *Conference Series, vol. 1770(1)*.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (27), 2025
ISSN 2663 - 4023

Myroslav Lakhno
Ggraduate student
National University of Life and Environmental Sciences of Ukraine, Kiev, Ukraine
ORCID ID: 0000-0001-6979-6076
valss725@gmail.com

SYSTEMATIC ANALYSIS OF DIGITAL TRACES IN THE UNIVERSITY INFORMATION AND EDUCATIONAL SYSTEM

Abstract. Under the conditions of rapid development of digital technologies and the growth of cyber threats to information systems, which are widely used in many areas of human activity, the relevance of mastering information security (IS) competencies is becoming increasingly obvious. Training university students in the IS not only expands their professional competencies, but also plays a key role in the formation of qualified specialists capable of effectively solving new challenges in the field of cyber threats. Such specialists will have a proactive position and the ability to self-organize under the conditions of constantly changing cyber threats, which is especially important in business processes built on digital technologies. This paper proposes an algorithm for a decision support system (DSS), aimed at improving the quality of education and the level of security of the university's digital educational environment (DSE). The algorithm is based on the analysis of digital footprints (DF) of users and can be implemented in the model of intelligent assistant for the CSOS. The essence of the proposed approach is to use matrix factorization of the DSS of users, which allows more effective management and analysis of information about the activities of students and teachers in the digital space. One of the key advantages of this approach is its ability to solve the problem of developing the competency profile of students, especially in the field of IS. The specified algorithm contributes to a deeper understanding and mastery of the necessary skills, which, in turn, significantly increases the degree of security of the CSOS and computer systems of universities. In the context of the growing number of cyber threats and the increasing complexity of their manifestations, the proposed solutions help to ensure reliable protection of the educational infrastructure and contribute to the training of specialists ready for the challenges of modern digital world.

Keywords: digital educational environment; university; digital footprints; matrix factorization; machine learning; information security; competencies.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Shevchuk, B. (2023). Virtualne osvitnie seredovyshche zakladu vyshchoi osvity: realii ta perspektyvy. *Nauka i tekhnika sohodni, 14(28)*, 493–504.
2. Buinytska, O. P., Varchenko-Trotsenko, L. O., & Hrytseliak, B. I. (2020). Tsyfrovizatsiia zakladu vyshchoi osvity. *Osvitolohichnyi dyskurs: elektronne naukove fakhove vydannia, (28)*, 64–79.
3. Areshonkov, V. Yu. (2020). Tsyfrovizatsiia vyshchoi osvity: vyklyky ta vidpovidi. *Visnyk Natsionalnoi akademii pedahohichnykh nauk Ukrainy, 2(2)*, 1–6.
4. Pozo, J. I., Pérez Echeverría, M. P., Cabellos, B., & Sánchez, D. L. (2021). Teaching and learning in times of COVID-19: Uses of digital technologies during school lockdowns. *Frontiers in Psychology, 12, 656776.*
5. Mustapha, I., Van, N. T., Shahverdi, M., Qureshi, M. I., & Khan, N. (2021). Effectiveness of digital technology in education during COVID-19 pandemic. *A bibliometric analysis*, 136–154.
6. Vaccari, A., Calabrese, F., Liu, B., & Ratti, C. (2009, November). Towards the SocioScope: an information system for the study of social dynamics through digital traces. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 52–61.
7. Osborne, N., & Connelly, L. (2015). Managing your digital footprint: Possible implications for teaching and learning. In *Proceedings of the 2nd European Conference on Social Media ECSM*, 354–361.
8. Morze, N., Kuzminska, O., & Mazorchuk, M. (2019). Attitude to the digital learning environment in Ukrainian universities. In *ICT in Education, Research, and Industrial Applications, vol. 2393*, 53–67.
9. Owoc, M. L., Sawicka, A., & Weichbroth, P. (2019, August). Artificial intelligence technologies in education: benefits, challenges and strategies of implementation. In *IFIP International Workshop on Artificial Intelligence for Knowledge Management*, 37–58.

10. Ahmad, S. F., Alam, M. M., Rahmat, M. K., Mubarik, M. S., & Hyder, S. I. (2022). Academic and administrative role of artificial intelligence in education. *Sustainability, 14(3), 1101.*
11. Goodwin, A. L. (2020). Globalization, global mindsets and teacher education. *Action in Teacher Education, 42(1),* 6–18.
12. Ren, Z., Xin, Y., Ge, J., Zhao, Z., Liu, D., Ho, R. C., & Ho, C. S. (2021). Psychological impact of COVID-19 on college students after school reopening: a cross-sectional study based on machine learning. *Frontiers in Psychology, 12, 641806.*
13. Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments, 9(1), 11,* 2–19.
14. Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education, 44(1),* 130–160.
15. Van den Beemt, A., Groothuijsen, S., Ozkan, L., & Hendrix, W. (2023). Remote labs in higher engineering education: engaging students with active learning pedagogy. *Journal of Computing in Higher Education, 35(2),* 320–340.
16. Hasibuan, M., Reynolds, M., Male, S., & Hassan, G. M. (2023). Role of Theory in Analysing the dynamic of Self-regulated Learning process based on students' event logs data: A scoping review. *ASCILITE Publications*, 126–136.
17. Nevin, A. D., Chen, Y., Yang, S., & Quan-Haase, A. (2022). Key considerations in the interpretation of digital trace data. *The SAGE handbook of social media research methods*, 54–66.
18. Hüllmann, J. A. (2019). The construction of meaning through digital traces. *Proceedings of the Pre-ICIS*, 1–5.
19. Aarthi, M. (2021). Using Users Profiling to Identifying an Attacks. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12 (7),* 795–802.
20. Subramanian, K., & Subramanian, K. (2020). Introducing the Splunk Platform. *Practical Splunk Search Processing Language: A Guide for Mastering SPL Commands for Maximum Efficiency and Outcome*, 1–38.
21. Shu, K., Zhou, X., Wang, S., Zafarani, R., & Liu, H. (2019, August). The role of user profiles for fake news detection. In *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining*, 436–439.
22. Sen, I., Flöck, F., Weller, K., Weiß, B., & Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. *Public Opinion Quarterly, 85(S1),* 399–422.
23. Coulter, R., Han, Q. L., Pan, L., Zhang, J., & Xiang, Y. (2019). Data-driven cyber security in perspective – Intelligent traffic analysis. *IEEE transactions on cybernetics, 50(7),* 3081–3093.
24. Lei, C., Dai, H., Yu, Z., & Li, R. (2020). A service recommendation algorithm with the transfer learning based matrix factorization to improve cloud security. *Information Sciences, 513,* 98–111.
25. Zheng, X., Guan, M., Jia, X., Guo, L., & Luo, Y. (2022). A matrix factorization recommendation system-based local differential privacy for protecting users' sensitive data. *IEEE Transactions on Computational Social Systems, 10(3),* 1189–1198.
26. Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms: an overview. *Journal of physics: conference series, 1142.*
27. Fedoruk, P. I., Pikuliak, M. V., & Dutchak, M. S. (2010). Intelektualnyў mekhanizm pobudovy indyvidualnoї navchalnoї traiektoriї v adaptyvnykh systemakh dystantsiўnoho navchannia. *Shtuchnyi intelekt, 3,* 668–678.
28. Joy, J., & Renumol, V. G. (2020, December). Comparison of generic similarity measures in E-learning content recommender system in cold-start condition. In *2020 IEEE Bombay section signature conference (IBSSC),* 175–179.
29. Xia, P., Zhang, L., & Li, F. (2015). Learning similarity with cosine similarity ensemble. *Information sciences, 307,* 39–52.
30. Mana, S. C., & Sasipraba, T. (2021, March). Research on cosine similarity and pearson correlation based recommendation models. *In Journal of Physics*: *Conference Series, vol. 1770(1).*