Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

Штучний інтелект

УДК 004.8:37.013:004.056

АЛГОРИТМ ОЦІНЮВАННЯ ДОСТОВІРНОСТІ ВІДПОВІДЕЙ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ ПРИ СТВОРЕННІ НАВЧАЛЬНОГО КОНТЕНТУ

Н.О. Маслова[1,2], О.М. Любименко[2,3]
[1]Lviv State University of Life Safety, Lviv, Ukraine
[2]Donetsk National Technical University, Drohobych, Ukraine
[3]Lutsk National Technical University, Lutsk, Ukraine
ORCID https://orcid.org/0000-0002-9078-0973
ORCID https://orcid.org/0000-0002-5935-6891
E-mail: nataliia.maslova@donntu.edu.ua, olena.liubymenko@donntu.edu.ua

АНОТАЦІЯ

У роботі проаналізовано ризики, пов’язані з коректністю та достовірністю навчального контенту, створеного за допомогою інструментів штучного інтелекту. Інтелектуальні інструменти на основі штучного інтелекту сприяють автоматизації процесу розроблення інтерактивних навчальних матеріалів, підвищенню рівня персоналізації навчання та оптимізації аналізу результатів освітньої діяльності. Водночас впровадження технологій штучного інтелекту в освітнє середовище супроводжується появою нових цифрових ризиків, зокрема поширенням дезінформації та формуванням залежності від технологічних засобів. У цьому дослідженні проаналізовано ризики, пов'язані з правильністю та достовірністю освітнього контенту, створеного за допомогою інструментів штучного інтелекту. Запропоновано алгоритм оцінювання достовірності відповідей систем штучного інтелекту, що використовуються під час створення навчального контенту. Алгоритм ґрунтується на моделюванні процесу перевірки достовірності відповідей, згенерованих штучним інтелектом, передбачає поетапний аналіз генерованих результатів, обчислення показників точності та визначення рівня їх достовірності на основі порівняння з контрольними джерелами. Проведено експериментальну оцінку кількох інструментів штучного інтелекту з використанням тестових завдань, пов'язаних з темами інформаційної безпеки. Результати показали, що точність відповідей, згенерованих ChatGPT, досягала приблизно 90–95%, тоді як інші інструменти демонстрували нижчу надійність залежно від складності завдання. Запропонований алгоритм спрямований на зменшення ризиків поширення дезінформації та сприяє підвищення якості навчальних матеріалів, створених із використанням інтелектуальних систем.

Ключові слова: *алгоритм оцінювання достовірності, моделювання, штучний інтелект, навчальний контент, цифрові освітні платформи.*

Вступ

Штучний інтелект (ШІ) став невід’ємним елементом сучасного цифрового середовища та активно застосовується в різних сферах діяльності. Технології ШІ використовуються для фільтрації спаму, виявлення фішингових атак, аналізу користувацьких запитів і обробки природної мови. Крім того, алгоритми штучного інтелекту застосовуються для автоматичного створення субтитрів, аналізу відеоматеріалів, визначення оптимальних маршрутів із урахуванням дорожньої ситуації та покращення якості машинного перекладу шляхом навчання на нових масивах даних.

Стрімкий розвиток технологій штучного інтелекту сприяв удосконаленню цифрових асистентів та інструментів для створення навчального контенту і значно розширив сфери їх використання. У сфері освіти ШІ допомагає викладачам у підготовці навчальних матеріалів, автоматизуючи рутинні процеси та забезпечуючи персоналізацію навчання. Такі системи можуть генерувати тестові завдання, вправи та навчальні плани, а також адаптувати тексти відповідно до рівня підготовки студентів. Автоматизовані системи оцінювання здатні перевіряти письмові роботи, аналізувати відповіді студентів і оцінювати рівень знань, що дозволяє викладачам оперативніше надавати зворотний зв’язок.

Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

Застосування адаптивних освітніх платформ забезпечує індивідуалізацію навчального процесу, оскільки навчальний контент підлаштовується під потреби та рівень підготовки кожного студента. Системи рекомендацій пропонують персоналізовані теми для вивчення та додаткові вправи, а чат-боти можуть надавати пояснення навчального матеріалу у будь-який зручний час. Загалом використання технологій штучного інтелекту дозволяє скоротити час підготовки навчальних матеріалів і підвищити ефективність освітнього процесу.

Питання використання сучасних цифрових технологій у сфері освіти вже розглядалися в наукових дослідженнях. Зокрема, у роботі [1] досліджено можливості підвищення якості освіти шляхом застосування інтерактивних навчальних інструментів і створення освітнього контенту на основі цифрових платформ дистанційного навчання. У роботі [2] проаналізовано роль інформаційних технологій у навчальному процесі, особливо в умовах дистанційної освіти, а також питання інформаційної безпеки в системах управління навчанням.

Разом з тим, активне використання генеративних систем штучного інтелекту в освітньому середовищі супроводжується появою нових цифрових ризиків, пов’язаних із можливістю формування недостовірних або частково некоректних відповідей. Це створює потенційні загрози поширення дезінформації у навчальному контенті та зумовлює необхідність розроблення ефективних механізмів перевірки достовірності інформації, сформованої інтелектуальними системами.

Незважаючи на активний розвиток інструментів штучного інтелекту, питання алгоритмічного оцінювання достовірності відповідей, сформованих такими системами під час створення навчального контенту, залишається недостатньо дослідженим. Існуючі підходи здебільшого орієнтовані на загальне оцінювання якості інформації, але не забезпечують формалізованого механізму перевірки достовірності відповідей ШІ у навчальному середовищі.

Аналіз літературних даних та постановка проблеми

Використання інструментів на основі штучного інтелекту в освітньому процесі сприяє автоматизації навчальної діяльності, персоналізації навчання та підвищенню ефективності освітнього середовища. Такі технології дають змогу створювати інтерактивний навчальний контент, адаптувати навчальні матеріали до індивідуальних потреб студентів і оптимізувати процедури оцінювання результатів навчання. Водночас впровадження штучного інтелекту в освітню практику має як переваги, так і певні обмеження.

До основних переваг належить підвищення ефективності навчального процесу завдяки автоматизації низки завдань, зокрема створення інтерактивних матеріалів, адаптації контенту до індивідуальних особливостей студентів та оперативного оцінювання результатів їхньої діяльності. Наприклад, віртуальні асистенти ChatGPT та Gemini можуть забезпечувати швидку підтримку студентів, надаючи відповіді на запитання та пояснюючи складні теми. Інструменти підтримки навчального процесу, такі як Redmenta та Amperia, надають можливості для розроблення сучасних навчальних матеріалів, тоді як адаптивні освітні платформи, зокрема Socratic і Khanmigo, сприяють індивідуалізації навчання, регулюючи темп і рівень складності завдань відповідно до потреб кожного студента.

Разом із тим використання технологій штучного інтелекту супроводжується певними ризиками. Зокрема, існує ймовірність формування надмірної залежності студентів від технологічних інструментів, що може негативно впливати на розвиток їхнього самостійного мислення та навичок критичного аналізу. До інших потенційних проблем належать питання захисту конфіденційності даних, оскільки значна кількість цифрових платформ здійснює збір і обробку персональної інформації користувачів, а також ризики виникнення алгоритмічних упереджень, які можуть впливати на об’єктивність результатів навчання. Наприклад, використання платформ для генерації навчального контенту, таких як Quizlet AI Tutor або Curipod, може призводити до створення неточного або недостатньо якісного матеріалу у разі некоректного налаштування інструментів. Крім того, застосування сервісів для розроблення презентацій, зокрема Gamma та Canva, може спричиняти надмірну стандартизацію навчальних матеріалів, що обмежує творчий підхід до їх підготовки.

У таблиці 1 наведено класифікацію поширених інструментів на основі штучного інтелекту, які сприяють автоматизації освітніх процесів, забезпечують адаптацію навчання до рівня підготовки студентів та допомагають викладачам створювати персоналізований і тематично орієнтований навчальний контент.

Використання інструментів штучного інтелекту (ШІ) у навчальному процесі відкриває нові можливості для розвитку освітнього середовища, проте одночасно супроводжується низкою потенційних ризиків і викликів [3]. У наукових джерелах розглядаються різні аспекти та типи таких вразливостей.

Зокрема, у роботі [4] виокремлено три основні проблеми. Першою з них є залежність від технологій. Надмірне використання інструментів ШІ може сприяти зниженню рівня критичного мислення та формуванню залежності студентів від автоматизованих систем. У випадках, коли інформація подається

ISSN 2786-9024

Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

Табл. 1. Інструменти з ШІ для підтримки навчального процесу

Тип інструмента	Назва/сайт розробника
Віртуальний асистент	ChatGPT (https://chatgpt.com/g/g-jekajgZGe-insight), Gemini (https://gemini.google.com), Microsoft Copilot (https://copilot.microsoft.com/chats/A6cUi2FsZvuJsZbDWSKg2), Claude (https://www.anthropic.com)
Помічники навчального процесу	Redmenta (https://redmenta.com), Gios (https://gioschool.com/ua), Amperia (https://edpro.ua/amperia), TWEE(https://app.twee.com/auth/signin?utm_source=chatgpt.com)
Засоби інтерактивного навчання та адаптивні платформи	Socratic (by Google) (https://socratic.org), Khanmigo (by Khan Academy) (https://khanmigo.ai/), Querium (https://www.querium.com), Carnegie Learning MATHia (https://www.carnegielearning.com/solutions/math/mathia/), Century Tech (https://www.century.tech/)
Інструменти створення презентацій	Gamma (https://gamma.app), Microsoft Designer (https://designer.microsoft.com), Canva (https://www.canva.com), Kahoot! (https://kahoot.com/
Генератори навчального контенту	Quizlet AI Tutor, Q-Chat (https://quizlet.com/qchat), Curipod (https://curipod.com), MagicSchool.ai (https://www.magicschool.ai/), Knowji – (https://www.knowji.com), ScribeSense (https://www.scribesense.com),
Помічники написання та аналізу текстів	Reading Coach (https://coach.microsoft.com/uk-ua), Grammarly for Education (https://www.grammarly.com),Quillbot (http://quillbot.com), Writefull (http://www.writefull.com)
Освітні чат-боти та репетитори	Tutor AI (http://tutorai.me), Edmentum Exact Path (http://www.edmentum.com/products/exact-path), Cognii (http://www.cognii.com), Squirrel AI (http://squirrelai.com)

у спрощеному вигляді, користувачі можуть приділяти менше уваги самостійному аналізу та розв’язанню складних завдань. У дослідженні Stanford University зазначено, що близько 60% студентів віком від 17 до 25 років схильні сприймати відповіді, згенеровані ШІ, без їхньої критичної перевірки.

Другою проблемою є питання конфіденційності та безпеки даних. Багато застосунків на основі ШІ здійснюють збір значних обсягів інформації про користувачів, і недостатній рівень її захисту може призвести до витоків персональних даних та порушення приватності. Згідно з дослідженням IBM, на яке посилається джерело [4], у 2023 році близько 12% випадків кіберзлочинів в освітньому секторі були пов’язані з витоками персональних даних, що підкреслює необхідність посилення заходів інформаційної безпеки.

Третьою проблемою є упередженість алгоритмів. Алгоритми штучного інтелекту можуть містити вбудовані упередження, що здатні призводити до неточних або несправедливих результатів. Наприклад, автоматизовані системи оцінювання можуть демонструвати необ’єктивність через обмеження моделей або використання нерепрезентативних даних. Дослідження MIT Media Lab показало, що ШІ-системи можуть формувати різні результати для користувачів із різним соціальним або культурним бекграундом.

У дослідженні [5] також наголошується на нерівності доступу до технологій. Не всі студенти мають однакові можливості користування сучасними цифровими інструментами та стабільним доступом до мережі Інтернет. Наприклад, студенти з сільських або соціально вразливих регіонів можуть не мати необхідних технічних засобів або якісного інтернет-з’єднання, що сприяє формуванню цифрового розриву та посиленню освітньої нерівності.

Ще два важливі виклики розглянуто у роботі [6]. Перший із них пов’язаний з витратами на впровадження та підтримку технологій ШІ. Інтеграція таких систем у навчальний процес потребує значних фінансових ресурсів для придбання програмного забезпечення, модернізації технічної інфраструктури та підготовки персоналу. Недостатній рівень фінансування може стати суттєвою перешкодою для ефективного впровадження цих технологій у закладах освіти.

Другим аспектом є трансформація ролі викладача. Автоматизація окремих педагогічних завдань за допомогою інтелектуальних систем може призвести до зміни функцій викладачів або навіть скорочення їхньої кількості, що потенційно впливає на якість навчального процесу та рівень взаємодії зі студентами.

Фактори з сьомого по дев’ятий розглянуто у джерелі [7]. Першим із них є обмеженість креативності та гнучкості. Оскільки системи ШІ працюють на основі алгоритмів і попередньо визначених правил, їхня здатність адаптуватися до нестандартних ситуацій або пропонувати нові підходи до розв’язання проблем може бути обмеженою.

Т4. №6(38)'2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

Восьмим фактором є алгоритмічна дискримінація, яка може виникати через використання упереджених даних або недосконалих алгоритмів, що призводить до нерівного ставлення до різних груп користувачів.

Дев'ятим фактором визначено зниження рівня соціальних і емоційних компетентностей, що може бути пов'язано з надмірною довірою до рішень, запропонованих системами ШІ.

Десятим аспектом є проблема так званих «галюцинацій» штучного інтелекту, розглянута у огляді [8]. Автор, аналізуючи близько тридцяти наукових джерел, зазначає, що сучасні системи ШІ можуть генерувати неправдиву або недостовірну інформацію, яка при цьому виглядає переконливою. Використання таких даних у навчальному процесі може призвести до формування помилкових уявлень про навчальний матеріал та негативно вплинути на якість знань.

Таким чином, застосування технологій штучного інтелекту в освіті порушує не лише питання етики, конфіденційності та інформаційної безпеки, а пов'язане з проблемами недостатнього рівня цифрової компетентності користувачів і фрагментарності навчального забезпечення, що підкреслюється у дослідженні [9].

Джерела [10–13] слугували основою для узагальнення матеріалів наступного розділу. Зокрема, у звіті [12] досліджено потенційні ризики використання чат-ботів на основі штучного інтелекту в освітньому середовищі, серед яких особливо підкреслюється небезпека поширення дезінформації. Автори зазначають, що неточні або хибні дані, згенеровані системами ШІ, можуть вводити студентів в оману, спотворювати навчальний контент і негативно впливати на результати навчання. У зв'язку з цим пропонується застосовувати збалансований підхід до інтеграції технологій ШІ, поєднуючи їх із традиційними методами навчання.

У дослідженні [13] проведено порівняння ефективності підказок, згенерованих системою ChatGPT, із підказками, створеними людськими репетиторами. Отримані результати свідчать, що хоча обидва типи підказок сприяють покращенню результатів навчання, рекомендації, надані викладачами, виявилися більш ефективними. Це підкреслює наявні обмеження навчального контенту, створеного за допомогою ШІ, а також важливість людського контролю для забезпечення якості та достовірності інформації.

Мета та задачі дослідження

Метою статті є розроблення алгоритму оцінювання достовірності відповідей систем штучного інтелекту, що використовуються при створенні навчального контенту.

Для досягнення мети дослідження потрібно виконати такі **завдання**:

1. Проаналізувати сучасні підходи до використання систем штучного інтелекту у створенні навчального контенту та визначити основні переваги і ризики їх застосування в освітньому середовищі.
2. Дослідити проблему достовірності інформації, сформованої генеративними системами штучного інтелекту, зокрема явище так званих «галюцинацій» моделей.
3. Визначити критерії оцінювання достовірності відповідей, сформованих системами штучного інтелекту під час підготовки навчальних матеріалів.
4. Розробити алгоритм оцінювання достовірності відповідей систем штучного інтелекту, який може бути використаний під час створення та перевірки навчального контенту.
5. Проаналізувати можливості практичного застосування запропонованого алгоритму у процесі підготовки навчальних матеріалів та оцінити його ефективність.

Матеріали та методи досліджень

На основі проведеного аналізу літературних джерел, а також з урахуванням практичного досвіду авторів щодо використання інструментів штучного інтелекту для створення навчального контенту було сформовано таблицю 2.

Інструменти, представлені в таблиці, були навмисно відібрані з урахуванням різних типів, відповідно до раніше запропонованої класифікації. Такий підхід дозволив охопити різні категорії інтелектуальних сервісів та проаналізувати характерні для них ризики.

Для обґрунтування інформації, наведеної у третій колонці таблиці, розглянемо окремі приклади. Зокрема, шаблони платформи Canva можуть містити елементи, що охороняються авторським правом, такі як зображення, ілюстрації або шрифти. У разі використання цих матеріалів без належних ліцензій чи дозволів, особливо при створенні комерційного контенту, існує ризик порушення авторських прав їхніх власників. Так, у 2021 році компанія Getty Images подала позов проти компанії Canva, звинувативши її у порушенні авторських прав щодо частини зображень, доступних на платформі.

Подібно до інших онлайн-сервісів для створення та зберігання цифрового контенту, Canva здійснює збір і зберігання персональних даних користувачів. До таких даних належать реєстраційна інформація, платіжні реквізити, а також створений користувачами контент (зображення, графіка тощо). Тому питання захисту персональної інформації також розглядається серед потенційних ризиків використання цієї платформи. Хоча протягом останніх років Canva не повідомляла про значні витоки даних, як і у випадку з будь-якими великими онлайн-системами, ризики,

ISSN 2786-9024

Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

Табл. 2. Вразливості інструментів створення навчального контенту з ШІ

Тип інструменту	Назва	Можливі вразливості
Віртуальний асистент	ChatGPT	Ненадійна інформація (галюцинації)
Помічники навчального процесу	Goodnotes	Безпека збереження нотаток, ризик втрати даних
Створення візуального контенту	Leonardo.ai	Авторське право на згенеровані зображення, використання шкідливого контенту, порушення етичних норм
Інтерактивне навчання	Khanmigo	Упередженість алгоритмів, неточність адаптації до учня (стереотипи, контент або відповіді, які не є нейтральними)
Створення презентацій	Canva	Використання шаблонів із потенційними авторськими обмеженнями, безпека особистих даних
Генерація навчального контенту	Quizlet AI Tutor	Неточність автоматично створених тестів, ризик плагіату
Помічники для написання та аналізу текстів	Grammarly for Education	Залежність від стабільності онлайн-доступу; застосування текстових даних користувачів для покращення алгоритмів (загроза витоку конфіденційної інформації)
Освітні чат-боти	Tutor AI	Відсутність контролю якості відповідей, ризик поширення дезінформації

пов’язані зі зберіганням персональної інформації, повністю виключити неможливо.

Ще одним прикладом є система Quizlet AI Tutor, яка може генерувати запитання, подібні до тих, що вже використовуються в інших навчальних матеріалах, підручниках або публікаціях. У ході експериментальної роботи було встановлено, що з 50 згенерованих системою запитань дві пари мали значну змістову подібність.

Певні питання щодо захисту даних виникали також у контексті використання сервісу Grammarly. Зокрема, у 2017 році компанія повідомила, що зберігає текстові дані користувачів з метою вдосконалення власних алгоритмів. Хоча розробники зазначали, що персональна інформація не використовується без згоди користувачів, така практика викликала дискусії щодо належного рівня захисту даних. Крім того, компанія зазнала критики через те, що зберігала введені тексти навіть у випадках, коли вони не були збережені користувачем.

У таблиці 3 наведено перелік загроз, пов’язаних із виявленими вразливостями, зазначеними у таблиці 2. Запропонована система загроз може бути використана як основа для оцінювання ризиків застосування інструментів штучного інтелекту в освітньому процесі та отримання подальших практичних результатів дослідження.

Слід також відзначити наукову роботу [14], яка була використана під час розроблення описаної нижче методики зниження ризику дезінформації. У зазначеному дослідженні запропоновано ефективний підхід до оцінювання діалогових систем, що може бути адаптований для аналізу точності та достовірності відповідей, сформованих віртуальними асистентами та освітніми чат-ботами.

Розвиваючи зазначену ідею, сформулюємо узагальнені рекомендації щодо зменшення ризиків і вразливостей, які можуть виникати під час створення навчального контенту з використанням інструментів на основі штучного інтелекту.

1. Віртуальні асистенти (ChatGPT, Google Bard тощо):

– обмежувати доступ до конфіденційної інформації, не вводити персональні або чутливі дані;

– здійснювати перевірку достовірності отриманих відповідей, використовуючи додаткові джерела для підтвердження фактів;

– контролювати використання таких інструментів у навчальному процесі та уникати повної автоматизації написання навчальних робіт.

2. Помічники організації навчального процесу (GoodNotes, Notion AI тощо):

– використовувати хмарні сервіси разом із локальним збереженням даних та регулярно здійснювати резервне копіювання;

– застосовувати захист доступу за допомогою паролів і двофакторної автентифікації;

– використовувати механізми шифрування інформації;

– контролювати рівень спільного доступу до документів.

3. Інструменти для створення візуального контенту (Leonardo.ai та ін.):

– перевіряти ліцензійні умови використання згенерованих зображень;

– застосовувати безпечні параметри генерації та фільтрувати небажаний контент;

– встановлювати правила використання таких інструментів і здійснювати контроль їх застосування в навчальних закладах.

4. Платформи інтерактивного навчання (Khanmigo, Duolingo AI тощо):

– здійснювати перевірку якості навчальних матеріалів викладачами;

Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

Табл. 3. Загрози інструментів створення навчального контенту з ШІ

Тип інструменту	Назва	Можливі загрози
Віртуальний асистент	ChatGPT	Дезінформація, витік особистих або конфіденційних даних
Помічники навчального процесу	Goodnotes	Несанкціонований доступ до нотаток, втрата даних через технічні збої
Створення візуального контенту	Leonardo.ai	Генерація маніпулятивних або шкідливих зображень, порушення авторських прав
Інтерактивне навчання	Khanmigo	Неправильне персоналізоване навчання, упередженість у рекомендаціях
Створення презентацій	Canva	Використання нелегального контенту, ризик фішингових атак при спільному доступі
Генерація навчального контенту	Quizlet AI Tutor	Автоматична генерація помилкового або застарілого матеріалу, шахрайство студентів
Помічники для написання та аналізу текстів	Grammarly for Education	Витік текстів користувачів, нав’язування шаблонного стилю
Освітні чат-боти	Tutor AI	Ненадійні або недостовірні відповіді, маніпуляція користувачами

– аналізувати алгоритми персоналізації з метою виявлення можливих упереджень;

– уникати надмірної залежності від ШІ та поєднувати його використання з традиційними методами навчання.

5. Інструменти для створення презентацій (Canva, Prezi AI тощо):

– перевіряти джерела використаних матеріалів та уникати ресурсів із невизначеним авторським правом;

– обмежувати публічний доступ до матеріалів без належного захисту;

– не завантажувати внутрішні або конфіденційні документи.

6. Генератори навчального контенту (Quizlet AI Tutor, Brisk Teaching тощо):

– здійснювати ручну перевірку згенерованих тестових матеріалів для запобігання появі неточних або некоректних запитань;

– обмежувати автоматичне оцінювання та використовувати ШІ як допоміжний інструмент;

– контролювати коректність рекомендацій і аналізувати отримані результати.

7. Інструменти для підтримки письма (Grammarly for Education, QuillBot тощо):

– контролювати передачу персональних даних і уникати введення конфіденційної інформації;

– запобігати надмірному спрощенню або зміні авторського стилю шляхом ручної перевірки текстів;

– не допускати автоматичного редагування студентських робіт без участі викладача.

8. Освітні чат-боти (Tutor AI, Squirrel AI тощо):

– здійснювати перевірку достовірності та коректності згенерованих відповідей і проводити їх регулярний моніторинг;

– використовувати фільтри для обмеження токсичного або небажаного контенту та налаштовувати безпечні параметри діалогу;

– регулярно оновлювати алгоритми роботи систем.

З огляду на викладене запропоновано алгоритм оцінювання достовірності відповідей систем штучного інтелекту, що використовується під час створення навчального контенту. Алгоритм застосовує методи моделювання й поетапний аналіз відповідей ШІ, обчислення показників точності та визначення рівня достовірності отриманих результатів.

Алгоритм оцінювання достовірності відповідей ШІ включає п’ять основних етапів.

Вхідні дані: запит користувача Q, відповідь системи штучного інтелекту A, множина контрольних джерел S.

Крок 1. Отримати відповідь A системи штучного інтелекту на запит Q.

Крок 2. Виконати порівняння відповіді A з інформацією з множини контрольних джерел S.

Крок 3. Обчислити показник точності:

$$Accuracy = \frac{N_{correct}}{N_{total}} \cdot 100\%,$$

де $N_{correct}$ – кількість правильних відповідей;

N_{total} – загальна кількість перевірених відповідей.

Якщо точність нижча за 90%, потрібно розробити додаткові механізми перевірки.

Крок 4. Визначити рейтинг достовірності відповіді за шкалою оцінювання.

Кожна відповідь ШІ може оцінюватися за шкалою достовірності:

1–3 бали – можлива дезінформація, потребує ретельної перевірки.

4–6 балів – частково правильна інформація, бажано перевірити додатково.

ISSN 2786-9024

Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

7–10 балів – висока достовірність, підтверджено джерелами.

Формула середньої достовірності:

$$Score = \frac{\sum_{i=1}^{N} Score_i}{N},$$

де $Score_i$ – оцінка достовірності кожної відповіді;

N – кількість оцінених відповідей.

Крок 5. Автоматичне позначення ненадійних відповідей. При цьому слід класифікувати відповідь як:

– ненадійну;

– частково достовірну;

– достовірну.

Якщо точність відповіді < 90% або середня достовірність < 7, система автоматично попереджає користувача про можливу дезінформацію та пропонує альтернативні джерела.

Крок 6. Впровадження комбінованої перевірки

За необхідності сформувати рекомендації щодо додаткової перевірки інформації.:

− алгоритмічна перевірка (перехресне порівняння відповіді ШІ з іншими джерелами);

− ручна перевірка експертом;

− оцінювання користувачами (рейтинги відповідей у системі).

Таким чином, запропонований алгоритм забезпечує комплексне оцінювання достовірності відповідей систем штучного інтелекту шляхом поєднання моделювання, автоматичного аналізу, експертної перевірки та користувацького оцінювання.

Наукова новизна запропонованого алгоритму полягає у комбінованому підході до оцінювання достовірності, що включає:

− автоматизоване оцінювання точності за допомогою розрахунку відсотка правильних відповідей та рейтингової шкали достовірності (більшість існуючих підходів використовують лише порівняння з референтними відповідями);

− динамічне ранжування надійності відповідей шляхом аналізу достовірності за шкалою 1–10 (більшість методів або оцінюють тільки загальну точність, або не використовують гнучкі рівневі підходи);

− адаптивне обмеження використання ненадійних відповідей – якщо достовірність < 7 або точність < 90%, система автоматично позначає відповідь як ненадійну та пропонує альтернативні джерела (в аналогічних дослідженнях (наприклад, [14]) оцінка обмежується лише порівнянням з контрольними відповідями, без активного впливу на користувацький досвід);

− інтегрований підхід (комбінована перевірка: автоматична + ручна + користувацькі оцінки).

Результати досліджень

Цей метод дозволяє гнучко аналізувати результати роботи ШІ, мінімізувати застосування невірних відповідей та підвищити якість використання ШІ при підготовці навчальних матеріалів й в освіті в цілому, дозволяє зменшити поширення недостовірної інформації у навчальному процесі.

Для проведення практичного експерименту було підібрано 10 тем з популярного й важливого напрямку – інформаційної безпеки. Були сформульовані завдання п’ятьох типів: а) поясни поняття; б) зроби розрахунок; в) виконай порівняння; г) створи запитання; д) надай відповідь. Завдання подавалося системам 5 разів з вимогою перевірити та уточнити результат. На останньому етапі порівнювалися відповіді й результати розрахунків, надані різними інструментами й класичні (вірні) відповіді. Приклад «функція – виклик – результат» наведено на рис.1.

Графічне відображення результатів експериментів, зокрема, для ChatGPT наведено на рис.2.

Обговорення результатів

Отже, для системи ChatGPT було отримано такі результати:

– середній рівень точності становив приблизно 90–95%;

– середній показник достовірності – 7–9 балів;

кількість відповідей, які були класифіковані як надійні, – 174;

– кількість ненадійних відповідей – 326.

Для системи Amperia – AI-асистента, призначеного для підтримки навчання, розв’язання задач і надання пояснень – середня точність становила приблизно 77,88%, а середній показник достовірності – близько 5,91. Із 500 проведених експериментів було отримано 124 надійні та 376 ненадійних відповідей.

Порівняльний аналіз точності та достовірності інструментів Redmenta і Gios, що використовуються як платформи для тестування та оцінювання, показав, що Redmenta демонструє вищу точність під час перевірки тестових завдань, особливо з варіантами закритих відповідей. Водночас система не завжди коректно оцінює розгорнуті відповіді. Платформа Gios характеризується нижчим рівнем точності, що може бути пов’язано з наявністю неоднозначних формулювань у тестових завданнях та залежністю результатів оцінювання від якості навчальних курсів. За результатами проведених розрахунків доцільно рекомендувати Redmenta для автоматизованого оцінювання тестів, тоді як Gios більш придатна для організації гнучкого навчального процесу та проведення навчального тестування.

Подібний аналіз було виконано і для платформ Quizlet AI Tutor (Q-Chat), Curipod, MagicSchool.ai, Knowji та ScribeSense, які використовуються для генерації навчального контенту. Найвищі показники точності продемонструвала система Knowji

Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

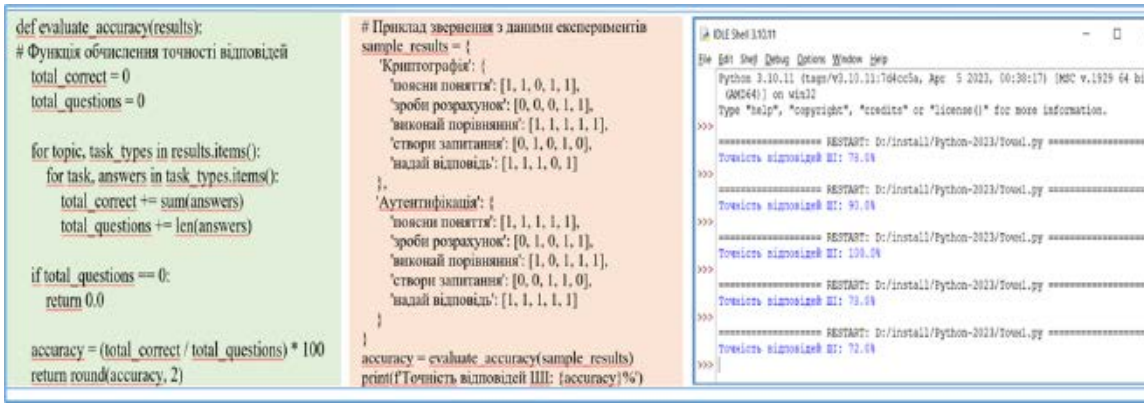

Рис. 1. Оцінювання точності відповідей ШІ

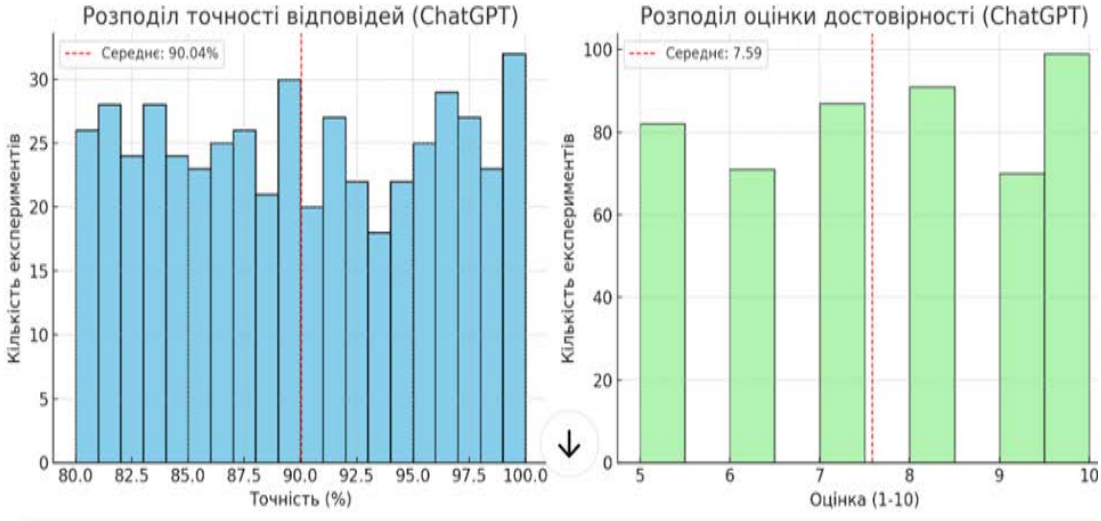

Рис. 2. Результати експериментів

(приблизно 85–95%), що, ймовірно, пов’язано з використанням структурованих даних та розширеного словникового ресурсу. Платформи Quizlet AI Tutor і MagicSchool.ai також показали високі результати, проте їхня точність значною мірою залежить від контексту поставлених завдань. Сервіси Curipod та ScribeSense продемонстрували нижчі показники, що, ймовірно, обумовлено залежністю від якості вихідних даних та складністю обробки текстової інформації.

Слід зазначити, що метою проведених експериментів не було визначення «ідеального» інструменту. Цифрові сервіси постійно вдосконалюються, проходять процеси навчання та регулярного оновлення. Розробники таких систем активно відстежують тенденції ринку та впроваджують новітні технології, алгоритмічні підходи й архітектурні рішення. На отримані результати впливає значна кількість чинників, зокрема мета дослідження, тип і якість використаних даних, їх обсяг, а також досвід користувача та рівень спеціалізації інструменту в певній предметній галузі.

Отримані результати відображають поточний стан досліджуваної проблеми. Основною метою дослідження було продемонструвати ефективність запропонованої методики контролю рівня достовірності та коректності відповідей, сформованих інструментами штучного інтелекту. Запропонований підхід може спростити процес вибору відповідного програмного забезпечення серед широкого спектра сучасних цифрових засобів, що використовуються для підготовки навчальних матеріалів.

Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

Висновки

Сучасні інформаційні технології та цифрові освітні платформи значно розширюють можливості навчального процесу, забезпечуючи доступ до інтерактивного контенту та персоналізованих освітніх ресурсів. Вони сприяють безперервному доступу до навчальних матеріалів, а також покращують комунікацію та взаємодію між учасниками освітнього процесу.

Водночас використання цифрових навчальних систем, зокрема інструментів штучного інтелекту, супроводжується низкою ризиків, пов’язаних із поширенням недостовірної інформації, витоками персональних даних та іншими загрозами інформаційної безпеки. Це зумовлює необхідність розроблення ефективних підходів до перевірки достовірності навчального контенту, сформованого інтелектуальними системами.

У роботі запропоновано алгоритм оцінювання достовірності відповідей систем штучного інтелекту, що використовується під час створення навчального контенту. Алгоритм ґрунтується на моделюванні процесу перевірки відповідей ШІ, обчисленні показників точності та визначенні рівня достовірності отриманих результатів. Його застосування дає змогу зменшити ризики поширення дезінформації та підвищити якість навчальних матеріалів, створених із використанням інструментів штучного інтелекту.

Запропонований алгоритм може бути використаний для вибору та оцінювання цифрових освітніх інструментів, а також для підвищення надійності навчального контенту в сучасному освітньому середовищі..

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно цього дослідження, у тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов’язаних даних.

ЛІТЕРАТУРА

[1] Г. Ю. Журавель та Н. О. Маслова, «Застосування інтерактивних засобів при створенні учбового контенту закладу освіти на основі цифрових навчальних платформ», у *Моделювання і комп’ютерна графіка: зб. матер. Восьмої міжнар. наук.-техн. конф.*, Донецьк: Донецький національний технічний університет, 2023, с. 95–100.

[2] Н. Маслова та О. Любименко, «Безпека та захист навчальних LMC систем», *Наукові праці Донецького національного технічного університету. Серія: Обчислювальна техніка та автоматизація*, т. 1, № 33, с. 38–46, 2023. doi: 10.31474/2786-9024/v1i1(33).299636.

[3] Khmelnitsky National University, *Proceedings of the Khmelnitsky National University. Technical sciences*. [Online]. Available: https://cpp.khmnu.edu.ua/index.php/cpp/issue/view/8

[4] S. Baumer, A. Carnevale, T. Corbett та C. Dumaresq, «Використання штучного інтелекту у навчанні: можливості та ризики», Вінницький державний педагогічний університет імені Михайла Коцюбинського, Buki, 30 бер. 2023. [Online]. Available: https://buki.com.ua/blogs/vikoristannia-stucnogo-intelektu-u-navcanni-mozlivosti-ta-riziki/

[5] NAUROK, «Методичні рекомендації: можливості, застереження та перспективи застосування ШІ на уроках української мови та літератури», 2023. [Online]. Available: https://naurok.com.ua/metodichni-rekomendaci-mozhlivosti-zasterezhennya-ta-perspektivi-zastosuvannya-shi-na-urokah-ukra-nsko-movi-ta-literaturi-450719.html (accessed Mar. 02, 2026).

[6] A. Kolomiiets та O. Kushnir, «Використання штучного інтелекту в освітній та науковій діяльності: можливості та виклики», *Modern Information Technologies and Innovation Methodologies of Education in Professional Training: Methodology, Theory, Experience, Problems*, вип. 70, с. 45–57, 2024. doi: 10.31652/2412-1142-2023-70-45-57.

[7] Н. С. Бобро, «Застосування штучного інтелекту у закладах вищої освіти: зарубіжний досвід», Noolab, 9 жовт. 2024. [Online]. Available: https://www.noolab.ch/ua/ua-blog/zastosuvannya-shtuchnogo-intelektu-u-zakladah-vishchoyi-osvity-zarubizhniy-dosvid

[8] Н. Баловсяк, «ШІ та виклики в освіті: як поєднати інноваційну технологію з консервативною традицією», Kunsht, 27 груд. 2024. [Online]. Available: https://kunsht.com.ua/articles/shi-ta-vyklyky-v-osviti-iak-poyednaty-innovatsiynu-tekhnolohiiu-z-konservatyvnoiu-tradytsiyeiu

[9] Інститут інформаційних технологій і засобів навчання НАПН України, *Використання штучного інтелекту в освіті*, 2024. [Online]. Available: https://lib.iitta.gov.ua/id/eprint/743864

[10] Міністерство освіти і науки України, *Інструктивно-методичні рекомендації щодо використання штучного інтелекту в закладах загальної середньої освіти*, Київ, 2024, с. 1–10. [Online]. Available: https://mon.gov.ua/static-objects/mon/sites/1/news/2024/05/21/Instruktyvno.metodychni.rekomendatsiyi.shchodo.SHI.v.ZZSO-22.05.2024.pdf (accessed Mar. 02, 2026).

[11] О. В. Кузьменко та К. Г. Гриценко, «Технології доброчесного використання штучного інтелекту у навчальному процесі», у *Матеріали конференції*, Центр українсько-європейських студій, 2024, с. 45–50. [Online].

Т4. №6(38)’2026 https://doi.org/10.32782/2786-9024/v4i6(38).359331

Available: https://cuesc.org.ua/images/informlist/Maket_advanced_training_PSAU.pdf

[12] S. Saghiri та A. Saghiri, *Catastrophic risks of AI-based chatbots in educational systems*. Society of Actuaries, 2024, с. 1–12. [Online]. Available: https://www.soa.org/4a3f4c/globalassets/assets/files/resources/research-report/2024/ai-risk-essays/saghiri-ai-based-chatbots.pdf

[13] Z. A. Pardos та S. Bhandari, «Learning gain differences between ChatGPT and human tutor generated algebra hints», *arXiv preprint arXiv:2302.06871*, 2023. [Online]. Available: https://arxiv.org/abs/2302.06871

[14] J. Deriu *et al.*, «Spot The Bot: A Robust and Efficient Framework for the Evaluation of Conversational Dialogue Systems», *arXiv preprint arXiv:2010.02140*, 2020. [Online]. Available: https://arxiv.org/abs/2010.02140

ALGORITHM FOR ASSESSING THE RELIABILITY OF ARTIFICIAL INTELLIGENCE SYSTEM RESPONSES IN EDUCATIONAL CONTENT CREATION

Nataliia Maslova, Olena Lyubymenko

The paper analyzes the risks associated with the correctness and reliability of educational content created using artificial intelligence tools. Intelligent tools based on artificial intelligence contribute to the automation of the process of developing interactive educational materials, increasing the level of personalization of learning and optimizing the analysis of educational results. At the same time, the introduction of artificial intelligence technologies into the educational environment is accompanied by the emergence of new digital risks, in particular, the spread of disinformation and the formation of dependence on technological means. This study analyzes the risks associated with the correctness and reliability of educational content created using artificial intelligence tools. An algorithm for assessing the reliability of responses of artificial intelligence systems used in the creation of educational content is proposed. The algorithm is based on modeling the process of checking the reliability of answers generated by artificial intelligence, provides for a step-by-step analysis of the generated results, calculation of accuracy indicators and determination of their reliability level based on comparison with control sources. An experimental evaluation of several artificial intelligence tools was carried out using test tasks related to information security topics. The results showed that the accuracy of answers generated by ChatGPT reached approximately 90–95%, while other tools demonstrated lower reliability depending on the complexity of the task. The proposed algorithm is aimed at reducing the risks of spreading disinformation and contributes to improving the quality of educational materials created using intelligent systems.

Keywords: *reliability assessment algorithm, modeling, artificial intelligence, educational content, digital learning platforms.*

REFERENCES

[1] H. Yu. Zhuravel and N. O. Maslova, “Application of interactive tools in creating educational content of educational institutions based on digital learning platforms” [“Zastosuvannia interaktyvnykh zasobiv pry stvorenni uchbovoho kontentu zakladu osvity na osnovi tsyfrovykh navchalnykh platform”], in *Modeling and Computer Graphics: Proc. 8th Int. Sci. and Tech. Conf.*, Donetsk National Technical University, 2023, pp. 95–100.

[2] N. Maslova and O. Liubymenko, “Security and protection of educational LMC systems” [“Bezpeka ta zakhyst navchalnykh LMC system”], *Scientific Works of Donetsk National Technical University. Series: Computing Technology and Automation*, no. 1(33), pp. 38–46, 2023. [Online]. Available: https://doi.org/10.31474/2786-9024/v1i1(33).299636

[3] Khmelnitsky National University, *Proceedings of the Khmelnitsky National University. Technical Sciences*. [Online]. Available: https://cpp.khmnu.edu.ua/index.php/cpp/issue/view/8

[4] S. Baumer, A. Carnevale, T. Corbett, and C. Dumaresq, “Use of artificial intelligence in education: opportunities and risks” [“Vykorystannia shtuchnoho intelektu u navchanni: mozhlyvosti ta ryzyky”], Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, Buki, 2023. [Online]. Available: https://buki.com.ua/blogs/vikoristannia-stucnogo-intelektu-u-navcanni-mozhlivosti-ta-riziki

[5] NAUROK, “Methodological recommendations: opportunities, cautions and prospects of AI use in Ukrainian language and literature lessons” [“Metodychni rekomendatsii: mozhlyvosti, zasterezhennia ta perspektyvy zastosuvannia ShI na urokakh ukrainskoi movy ta literatury”], 2023. [Online]. Available: https://naurok.com.ua/metodichni-rekomendaci-mozhlivosti-zasterezhennya-ta-perspektivi-zastosuvannya-shi-na-urokah-ukra-nsko-movi-ta-literaturi-450719.html (accessed Mar. 02, 2026).

[6] A. Kolomiiets and O. Kushnir, “Use of artificial intelligence in educational and scientific activities: opportunities and challenges” [“Vykorystannia shtuchnoho intelektu v osvitnii ta naukovii diialnosti: mozhlyvosti ta vyklyky”], *Modern Information Technologies and Innovation Methodologies of Education in Professional Training: Methodology, Theory, Experience, Problems*, vol. 70, pp. 45–57, 2024. [Online]. Available: https://doi.org/10.31652/2412-1142-2023-70-45-57

[7] N. S. Bobro, “Application of artificial intelligence in higher education institutions: foreign experience” [“Zastosuvannia shtuchnoho intelektu u zakladakh vyshchoi osvity: zarubizhnyi dosvid”], Noolab, 2024. [Online]. Available: https://www.noolab.ch/ua/ua-blog/zastosuvannya-shtuchnogo-intelektu-u-zakladah-vishchoyi-osvity-zarubizhniy-dosvid

[8] N. Balovsiak, “AI and challenges in education: how to combine innovative technology with conservative tradition” [“ShI ta vyklyky v osviti: yak poiednaty innovatsiinu

tekhnolohiiu z konservatyvnoiu tradytsiieiu”], Kunsht, 2024. [Online]. Available: https://kunsht.com.ua/articles/shi-ta-vyklyky-v-osviti-iak-poyednaty-innovatsiynu-tekhnolohiiu-z-konservatyvnoiu-tradytsiyeiu

[9] Institute of Information Technologies and Learning Tools of the NAES of Ukraine, “Use of artificial intelligence in education” [“Vykorystannia shtuchnoho intelektu v osviti”], 2024. [Online]. Available: https://lib.iitta.gov.ua/id/eprint/743864

[10] Ministry of Education and Science of Ukraine, *Instructional and methodological recommendations on the use of artificial intelligence in general secondary education institutions* [“Instruktyvno-metodychni rekomendatsii shchodo vykorystannia shtuchnoho intelektu v zakladakh zahalnoi serednoi osvity”], pp. 1–10, 2024. [Online]. Available: https://mon.gov.ua (accessed Mar. 02, 2026).

[11] O. V. Kuzmenko and K. H. Hrytsenko, “Technologies of ethical use of artificial intelligence in the educational process” [“Tekhnolohii dobrochesnoho vykorystannia shtuchnoho intelektu u navchalnomu protsesi”], in *Conference Proceedings*, Center for Ukrainian-European Studies, 2024, pp. 45–50.

[12] S. Saghiri and A. Saghiri, “Catastrophic risks of AI-based chatbots in educational systems,” *Society of Actuaries Research Report*, pp. 1–12, 2024. [Online]. Available: https://www.soa.org/4a3f4c/globalassets/assets/files/resources/research-report/2024/ai-risk-essays/saghiri-ai-based-chatbots.pdf

[13] Z. A. Pardos and S. Bhandari, “Learning gain differences between ChatGPT and human tutor generated algebra hints,” *arXiv preprint*, pp. 1–12, 2023. [Online]. Available: https://arxiv.org/abs/2302.06871

[14] J. Deriu *et al.*, “Spot The Bot: A Robust and Efficient Framework for the Evaluation of Conversational Dialogue Systems,” *arXiv preprint arXiv:2010.02140*, 2020. [Online]. Available: https://arxiv.org/abs/2010.02140

Дата першого надходження статті до видання: 13.02.2026
Дата прийняття статті до друку після рецензування:10.03.2026
Дата публікації (оприлюднення) статті: 12.05.2026