

doi: 10.32620/oikit.2026.107.08

УДК 004.89

О. І. Феоктистова, Я. М. Альошин

Підхід до розширення функціональності штучного інтелекту у формі chat gpt із використанням технології pinecone

Національний аерокосмічний університет «Харківський авіаційний інститут»

У статті розглянуто проблему різноманітності веб інтерфейсів, що впливає на ефективність бізнес-аналізу, зокрема на часові витрати та пошук необхідної інформації. Метою дослідження є розроблення теоретико-методологічного підґрунтя для розширення функціональності систем штучного інтелекту шляхом використання технології векторної семантичної бази даних. Запропонований підхід спрямований на оптимізацію процесів пошуку, структурування та інтерпретації даних у процесах бізнес-аналізу, що підвищує продуктивність роботи бізнес-аналітиків і якість отриманих результатів.

Об'єктом дослідження є процеси бізнес-аналізу із застосуванням систем штучного інтелекту, а предметом — методичне й програмне забезпечення для їх використання на основі технології векторної семантичної бази даних. Основна ідея полягає у створенні комплексу методів і програмних рішень, які забезпечують ефективну взаємодію людини та машини й скорочують часові витрати під час виконання аналітичних операцій.

У ході дослідження здійснено критичний огляд сучасного стану проблеми використання мовних моделей у бізнес-аналітиці, удосконалено методи людино-машинної взаємодії та розроблено експериментальний прототип для перевірки ефективності запропонованих рішень. Результати експерименту підтвердили підвищення ефективності розв'язання прикладних завдань бізнес-аналізу за допомогою мовних моделей, розширеного технологією семантичного пошуку та співвідношення контенту.

Отримані результати можуть бути застосовані у суміжних сферах цифрової аналітики, обробки великих масивів даних, електронної комерції та управління інформаційними потоками в мережі Інтернет. Це підвищує універсальність і практичну цінність запропонованого підходу, роблячи його перспективним напрямом для подальшого розвитку інтелектуальних аналітичних систем, інтеграції їх із семантичними базами даних та підвищення ступеня їх адаптивності до виникнення динамічного веб середовища.

Ключові слова: ШІ, семантичний пошук, штучний інтелект, бізнес-аналіз, семантичні векторні бази даних, людино-машинна взаємодія, оптимізація даних.

Вступ

Сучасні процеси бізнес-аналізу характеризуються високим рівнем складності, що зумовлено значною різноманітністю веб інтерфейсів, неоднорідністю структури інформаційних ресурсів та постійним зростанням обсягів даних. Гетерогенність джерел даних, відмінності у форматах подання та способах навігації ускладнюють пошук релевантної інформації, її обробку та інтерпретацію. Це, своєю чергою, призводить до збільшення часових витрат аналітиків, виникнення помилок у ході інтерпретації результатів бізнес-аналізу та дублювання зусиль під час проведення досліджень або прийняття управлінських рішень.

За таких умов актуальним є створення інтелектуальних інструментів інструментальних засобів, за допомогою яких стають можливими уніфікований доступ до інформаційних ресурсів, автоматизація процесів відбору та аналізу даних, а також зниження людського фактора при здійсненні рутинних операцій. Особливої уваги заслуговує інтеграція систем штучного інтелекту (ШІ) із сучасними семантичними технологіями, що дозволяють забезпечити глибше

розуміння змісту інформації та підвищити якість аналітичних висновків.

У цьому контексті запропоновано підхід до розширення функціональності систем ШІ типу ChatGPT шляхом їх інтеграції з векторними семантичними базами даних, які забезпечують високоточне семантичне зіставлення запитів і контенту. Така інтеграція спрямована на підтримку повного циклу аналітичної обробки даних за моделлю «пошук → осмислення → відповідь», що дозволяє істотно скоротити час виконання аналітичних операцій, підвищити релевантність отриманих результатів і створити передумови для створення або розроблення формування інтелектуальних систем підтримки прийняття рішень нового покоління.

Аналіз останніх досліджень і публікацій засвідчує зростання з боку наукової спільноти до впровадження технологій штучного інтелекту (ШІ) в українському науковому просторі. Андрій Дубчак [4] і Ярослав Литвиненко [16] розглядають сучасні напрями використання ШІ в умовах цифрової трансформації, тоді як Юрій Перучок, Анастасія Банарцева та Лілія Малигіна аналізують його переваги й ризики у контексті лінгводидактичних систем [6]. У сфері освіти Володимир Грицишин, Андрій Шевченко та Ганна Андрощук [2] вивчають роль ШІ у модернізації педагогічних процесів, а Олег Хорошайло та Сергія Кочергіна [8] наголошують на його вплив на якість викладання іноземних мов. Значний внесок у розвиток штучного інтелекту зробила Яна Бромберг [20] — дослідниця у галузі біоінформатики, що застосовує алгоритми ШІ у прогнозуванні функцій білків, та академік НАН України Анатолій Морозов [18], який заклав методологічні засади сучасних інтелектуальних систем. Сукупність цих досліджень формує теоретичне підґрунтя для подальшого розвитку ШІ в Україні.

Мета дослідження полягає у підвищенні ефективності процесу обробки веб-ресурсів при реалізації завдань бізнес-аналізу за рахунок створення теоретико-методичного базису для розроблення прототипу програмного забезпечення, що поєднує великі мовні моделі (LLM) із технологією створення й підтримки в актуальному стані векторної семантичної бази даних. Об'єктом дослідження виступають процеси бізнес-аналізу із залученням систем штучного інтелекту. Безпосередньо предмет дослідження – методичне та програмне забезпечення, що надає змогу застосовувати LLM у бізнес-аналізі, на основі векторної семантичної індексації Pinecone.

Наукова новизна дослідження полягає у розробленні теоретико-методичного підходу до інтеграції систем штучного інтелекту з технологією векторних семантичних баз даних що забезпечує підвищення ефективності бізнес-аналітичних процесів. Вперше формалізовано чотирикроковий конвеєр людино-машинної взаємодії, що охоплює етапи відбору контенту, семантичного зіставлення «питання–контент», генерації уточнювальних запитів і агрегації відповідей. Новизна також полягає у поєднанні чинників векторної та діалогової релевантності в межах єдиної моделі з керованою реконструкцією контексту, що забезпечує підвищення точності аналітичних результатів. Крім того, запропоновано систему метрик оцінювання, яка враховує інформаційну точність, логічну узгодженість і часову ефективність у прикладних бізнес-сценаріях.

Виклад основного матеріалу

Бізнес-аналіз – це багаторівнева система методів, підходів і інструментів, спрямованих на виявлення проблем, можливостей і тенденцій розвитку підприємства, а також на обґрунтування управлінських рішень і впровадження

організаційних змін. У науковій площині бізнес-аналіз розглядається як міждисциплінарна діяльність, що поєднує знання з економіки, менеджменту, інформаційних технологій, маркетингу та аналітики даних. Його основна мета полягає у перетворенні інформації на практичні знання, здатні забезпечити стратегічну стійкість і конкурентні переваги підприємства.

Бізнес-аналіз здійснюється на трьох рівнях – стратегічному, тактичному й операційному, кожен із яких спрямований на забезпечення ефективності управління підприємством. Стратегічний рівень охоплює аналіз ринку та прогнозування тенденцій розвитку, тактичний – планування й контроль реалізації управлінських рішень, а операційний – оптимізацію бізнес-процесів і автоматизацію рутинних завдань. Роль бізнес-аналітика в умовах цифрової трансформації виходить за межі традиційних функцій збору даних і включає моделювання процесів, розробку рекомендацій і координацію комунікацій між зацікавленими сторонами. [21] Інтеграція систем штучного інтелекту у поєднанні з технологією векторних баз даних, значно розширює можливості бізнес-аналітики, забезпечуючи автоматизовану обробку неструктурованої інформації, виявлення семантичних зв'язків і формування релевантних висновків, що підвищує точність, швидкість та обґрунтованість управлінських рішень [1, с. 78].

Розроблений у ході дослідження підхід містить такі етапи

1. Автоматизований відбір контенту з використанням алгоритму, який забезпечує комплексне сканування веб ресурсу та побудову повного списку сторінок для подальшої семантичної індексації. Перша стадія цього етапу являє собою послідовність операцій, спрямованих на вилучення, очищення й уніфікацію даних. Конкретний зміст є таким: спочатку система приймає вхідний параметр — початкову URL-адресу ресурсу, після чого звертається до файлу robots.txt з метою виявлення доступних карт сайту. Далі “екстрактор” переходить за знайденим посиланням на sitemap_index.xml, де здійснюється парсинг усіх наявних URL-адрес, зазвичай у межах 10–100 записів. Кожен знайдений sitemap завантажується у форматі розширеної розмітки XML, після чого з нього видаляються службові теги, залишаючи лише чисті посилання на сторінки. Наступною стадією є нормалізація отриманих даних; усунення дублікатів, видалення неангломовних сторінок (у базовій MVP-версії), а також формування єдиного інтегрованого списку уніфікованих URL-адрес. Цей список слугує основою для подальшого процесу семантичної обробки й формування ембедингів [1, с. 22].

На наступній стадії здійснюється завантаження даних за допомогою спеціалізованого скрейпера/парсера, який забезпечує структурну нормалізацію контенту з веб джерел. Алгоритм його роботи складається з трьох послідовних фаз:

I. Видалення HTML-шуму — усунення зайвих елементів, таких як рекламні блоки, скрипти, метатеги або дубльований текст, що не містять аналітичної цінності;

II. Уніфікація форматів і кодувань — приведення даних до єдиного стандарту (UTF-8, JSON або TXT) для коректної подальшої обробки;

III. Сегментація тексту на чанки розміром 300–800 слів із перекриттям 10–15 %, що дозволяє зберігати контекстну зв'язність і забезпечує оптимальну якість векторизації.

Далі формується керований перелік доменів і URL-адрес, який охоплює офіційні веб сайти, звіти, технічну документацію та репозиторії знань. Для

кожного текстового фрагмента обчислюється векторне подання $e_i \in R_d$ за допомогою сучасних моделей ембедингів, після чого результати зберігаються у векторній базі даних разом із метаданими (URL, заголовок, час створення, показники якості) [18, с. 120].

Таким чином, реалізований підхід забезпечує системність, точність і відтворюваність процесу первинного збору та структурування даних, створюючи оптимальні умови для подальшої семантичної індексації, пошуку й аналітичної обробки текстової інформації.

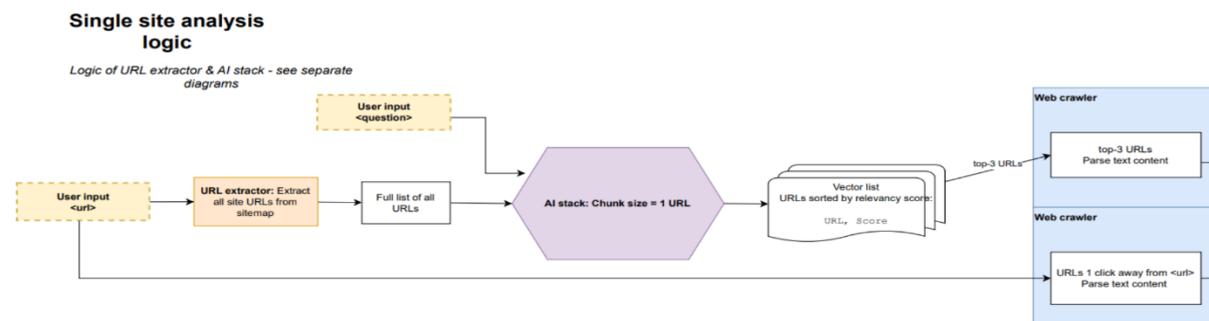


Рис. 1. Логіка співвідношення запиту користувача до контенту з використанням векторної бази даних. [12, с. 98]

2. Семантичне зіставлення питання до контенту. На другому етапі здійснюється зіставлення сформульованого запиту користувача з попередньо індексованими фрагментами контенту у векторному просторі. Для цього використовується SVDB (векторна, семантична база даних), у якій кожен елемент має векторне подання, що дає змогу визначати міру близькості між запитом і контентом за допомогою косинусної подібності. Алгоритм співвіднесення запиту користувача з релевантними частинами веб сайту (URL та метаданими) наведено на рис. 1 та 2.

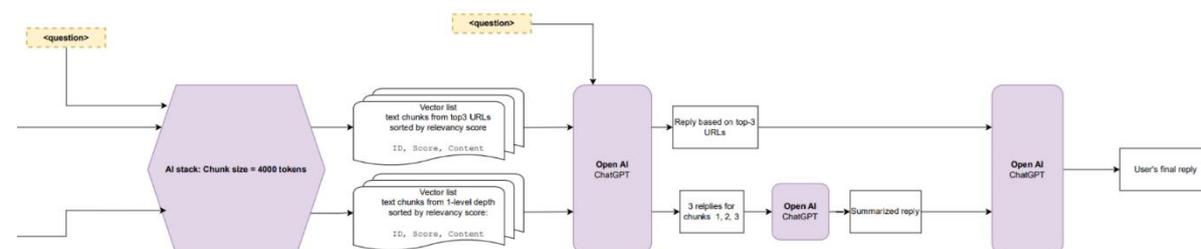


Рис. 2. Схема процесу побудови embedding-моделей та вибору релевантних сторінок для аналізу із використанням промптів

Система приймає вхідні дані у вигляді URL і текстового запиту, після чого AI-модуль формує векторне представлення запиту та здійснює пошук найбільш подібних фрагментів у базі. На основі отриманих результатів формується список найрелевантніших сторінок, які додатково аналізуються з використанням LLM для уточнення змістових зв'язків [5, с. 340]. Користувацьке питання q кодується у вектор e_q . Виконується пошук $top-k$ найближчих чанків найбільш релевантних до питання. Для підвищення надійності застосовується re-ranking (наприклад, cross-

encoder) і фільтри за метаданими (час, джерело, мова). Результатом є контекст $C_k = \{c_1, \dots, c_k\}$, що максимізує покриття й специфічність.

3. Формулювання запитів і генерація відповідей.

На етапі формулювання запитів до відібраного контенту (reasoning-oriented prompting) LLM (наприклад, ChatGPT) отримує керовану підказку з інструкціями щодо ролі, стилю відповіді, формату виходу та використання лише релевантних джерел (C_k). Застосовується ітеративна стратегія self-ask/query refinement, за якої модель уточнює запити до SVDB, розширюючи контекст і підвищуючи точність міркувань. Це забезпечує поєднання векторної та діалогової релевантності й прозору реконструкцію контексту. На етапі агрегації відповідей проміжні результати узгоджуються, усуваються дублювання та суперечності, а фінальна відповідь проходить перевірку достовірності на основі джерел C_k . Система може формувати структурований вихід (таблиці, JSON, маркери рішень), придатний для інтеграції у звітні чи аналітичні панелі [11, с. 15]. На основі C_k формується підказка для LLM із інструкціями: роль, тон, формат, цитування (за потреби). Застосовується iterative query formulation: модель генерує уточнювальні підзапити до векторної бази даних, доповнюючи контекст циклами «запит → отримання → узгодження релевантності». Це узгоджує ситуативний контекст користувача з документним контекстом джерел [15].

4. Агрегація та постобробка. Відповіді на кожній з ітерацій агрегуються зі зменшенням дублювання, відсіканням суперечливих фрагментів і коротким раціоналом. За потреби застосовується fact-checking через другий виклик LLM над цитованими уривками C_k .

Псевдокод конвеєра:

I. Input: q , конфігурація доменів і метрик.

II. $C_k \leftarrow \text{Pinecone.search}(eq, k, \text{filters})$

III. Repeat until convergence/limits:

$r \leftarrow \text{LLM}(q, C_k, \text{instructions})$; $q' \leftarrow \text{LLM.refine}(q, r)$;

$C_k \leftarrow C_k \cup \text{Pinecone.search}(eq', k')$

IV. $a \leftarrow \text{LLM.aggregate}(\{r_i\})$; post-processing.

V. Output: кінцева відповідь a + (опційно) посилання на джерела [2, с. 248].

На основі множини релевантних фрагментів формується керована підказка (prompt) для LLM. Для кожного фрагмента утворюється структурований блок: Питання: $[q]$; Контент: $[c_i]$. Підказка містить інструкції щодо ролі моделі (аналітик/експерт), стилю викладу (науковий, лаконічний), формату виходу (JSON/ієрархічний текст), політики цитування (посилання на c_i) та режим closed-book — використання лише наданого контексту C_k . З урахуванням ліміту контексту (≈ 4 – 8 тис. токенів) застосовуються пакетування чанків і пріоритизація за релевантністю (top- m mm, диверсифікація за доменом/датою). Використовується ітеративна стратегія self-ask/query refinement: модель генерує уточнювальні підзапити q до SVDB для непокритих аспектів і розширює контекст. Ітерації тривають до збіжності або досягнення обчислювальних лімітів. Для кожного пакета формується проміжна відповідь з обов'язковим зазначенням джерел (ID URL/фрагмента), що фіксує ланцюжок міркувань [8, с. 127].

Агрегація проміжних висновків і формування фінальної відповіді. Набір проміжних відповідей проходить поетапну агрегацію: [11, с. 15]. Узгодження змісту (consistency). Виконується виявлення дублювань і суперечностей між гі за допомогою евристик і повторного виклику LLM з інструкцією compare-and-contrast. Суперечливі твердження перевіряються через evidence check на вихідних

уривках Ск.

1. Злиття фактів (merge). Створюється узагальнена чернетка R з тематичною структурою (напр., Background – Findings – Implications або розгортка за критеріями бізнес-завдання). Зберігаються посилання на джерела рівня абзацу/твердження.

2. Оцінювання достовірності (faithfulness). Модель перевіряє, чи кожне твердження в R підтримується хоча б одним фрагментом із Ск (мета—мінімізувати галюцинації). За потреби виконується переформулювання із суворим обмеженням на джерела.

3. Фінальний виклад формується у науковому стилі з чіткою логіко-композиційною структурою та можливістю експорту у звітні формати (DOCX, JSON, Markdown). Для прикладних сценаріїв передбачено структурований вихід — таблиці показників, списки ризиків і рекомендації. З метою контролю якості ведеться журнал параметрів (версії ембедингів, пошукові пороги, кількість ітерацій, час обчислень), що забезпечує відтворюваність і порівнянність результатів. Оцінювання здійснюється за поєднанням інформаційних та процесних метрик — faithfulness, coverage, clarity/structure, time-to-answer та кількість ітерацій уточнення, що гарантує наукову обґрунтованість і практичну цінність отриманих висновків.

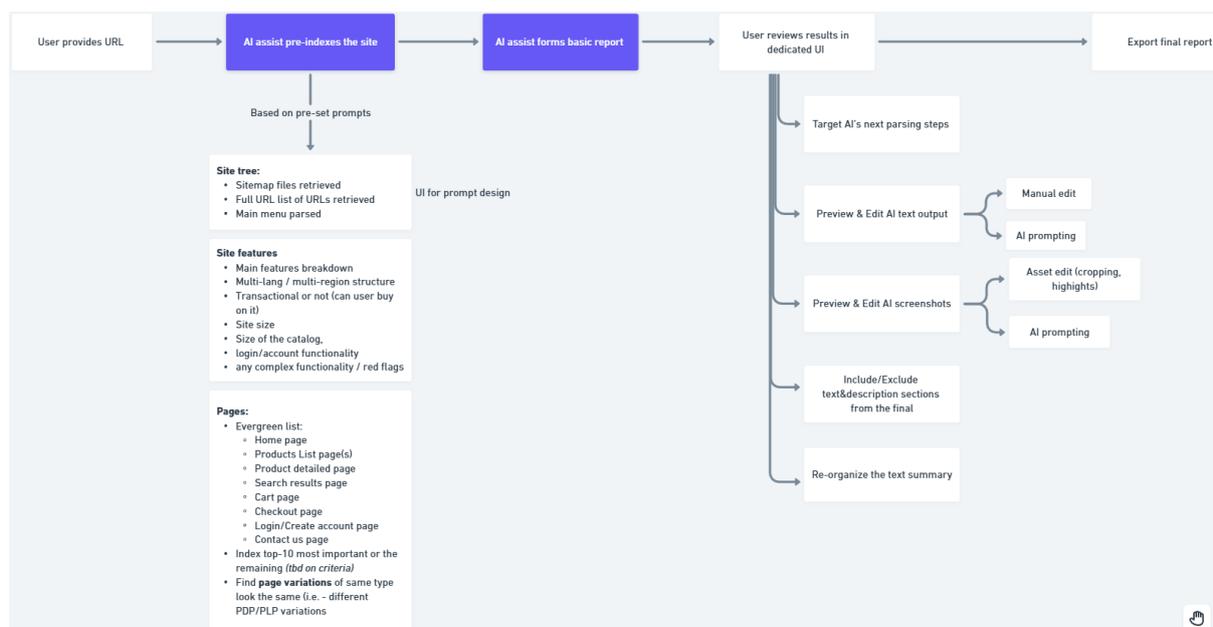


Рис. 3. Алгоритм попередньої індексації веб сайту системою та формування базового AI-звіту

Опис дослідницького прототипу (рис. 4) реалізує наведений конвеєр як веб-сервіс «Chat для бізнес-аналітики». Ключові компоненти: модуль завантаження та індексації джерел; векторно-семантичний індекс; LLM-оркестрація з контролем підказок; панель для аналізу трейсів запитів і порівняння відповідей. Забезпечено логування кроків для відтворюваності: версії ембедингів, параметри пошуку (k, пороги), тексти підказок, ітерації узгодження [14].

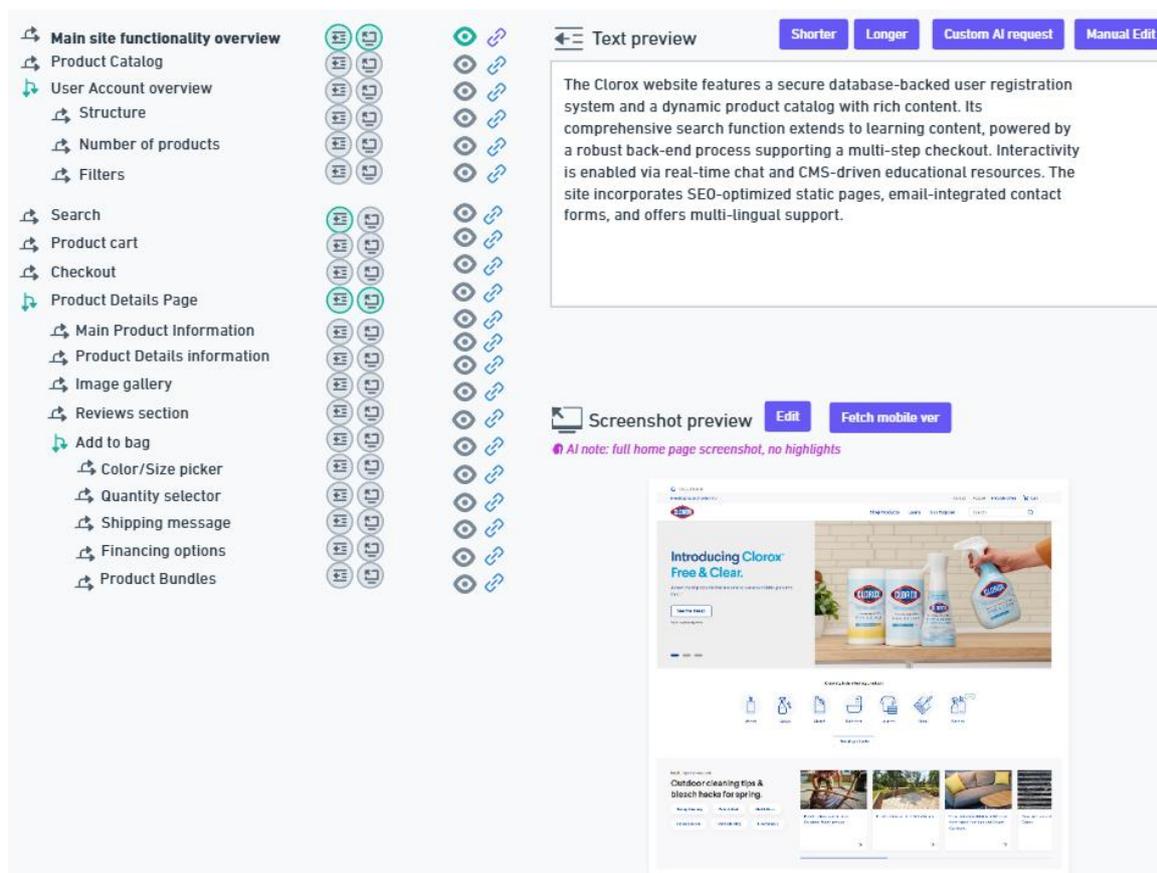


Рис. 3. Алгоритм попередньої індексації веб сайту системою та формування базового AI-звіту

Для оцінювання ефективності функціонування прототипу використано змішану методіку, що поєднує інформаційні та процесні метрики. До інформаційних метрик належать $Precision@k$ і $nDCG@k$ для оцінки якості ранжування, Answer Faithfulness — для визначення частки тверджень, підтверджених джерелами, та Coverage — для вимірювання повноти охоплення змісту запиту. Процесні показники включають TTA (time-to-answer) — час отримання релевантної відповіді, CTR-Refine — кількість ітерацій уточнення запиту, і Human Effort — суб'єктивну оцінку зменшення частки ручної роботи за шкалою Лайкерта [1, с. 78]. Порівнюються три умови: (А) базова робота аналітика (ручний пошук/читання); (В) класичний LLM без зовнішнього знання; (С) запропонований алгоритмом (конвеєр RAG з ітераціями). Вибірка завдань охоплює стратегічні, тактичні й операційні сценарії (напр., аналіз ринку, бенчмаркінг, витяг KPI з веб-звітів). Проводиться А/В-тестування із засліпленим оцінюванням якості відповідей експертами [19, с. 59].

Результати

Експерименти демонструють статистично значуще зниження TTA (time-to-answer), підвищення $Precision@k$ і faithfulness порівняно з умовами А і В. Інтеграція векторної семантичної бази даних забезпечує стабільну релевантність у динамічних веб-середовищах, а керована LLM-агрегація — когерентність і повноту відповідей. Для бізнес-аналітиків це трансформується у меншу фрагментацію робочих сесій, кращу відтворюваність і швидше формування

артефактів (довідки, порівняльні таблиці, короткі резюме).

Обговорення

Поєднання векторного пошуку з діалоговим міркуванням LLM мінімізує втрати смислу, спричинені різними UX-патернами (user-experience) сайтів, і робить аналіз менш залежним від особливостей навігації. Обмеження стосуються якості первинних джерел (шум, упередженість), залежності від вибору ембедингів і можливих галюцинацій LLM — тому важливими є прозоре цитування, журналювання кроків та політика human-in-the-loop. [3, с. 99].

Практична значущість запропонованого підходу полягає у можливості його масштабування в суміжні сфери: цифрову аналітику, обробку великих масивів даних, e-commerce, конкурентну розвідку, підтримку прийняття рішень. Також придатна до масштабування на різні домени (ринкові огляди, конкурентний аналіз, витяг KPI (key performance indicator), e-commerce), забезпечує контроль за джерелами знань і дає змогу формувати нормативно придатні звіти. Ключовою перевагою є зменшення часових витрат аналітиків при одночасному зростанні пояснюваності та відстежуваності джерел, що робить систему релевантною для корпоративних сценаріїв із високими вимогами до якості рішень. Він підвищує універсальність і зменшує бар'єри входу для нових аналітиків завдяки уніфікованому доступу до знань і відтворюваним процедурам. Розроблено методичний каркас і прототип RAG-системи для бізнес-аналізу, що інтегрує LLM із SVDB (спеціалізована векторна база даних). Підхід забезпечує кращу релевантність і швидкодію, ніж ручний пошук або «чистий» LLM, та підвищує адаптивність до динамічного веб середовища. Подальша робота спрямовується на: (i) автоматичний контроль достовірності (consistency checking), (ii) персоналізацію підстилі аналітика, (iii) багатомовну індексацію та (iv) приватність/безпеку корпоративних даних.

Керування якістю та відтворюваністю. У кожному вузлі конвеєра зберігаються журнали параметрів (версії ембедингів, пороги, k, текст підказок, кількість ітерацій), що забезпечує експериментальну відтворюваність. Якість вимірюється комбінованими метриками: інформаційними (Precision@k, nDCG@k, coverage, faithfulness) та процесними (time-to-answer, кількість ітерацій уточнення, частка ручних правок). Така організація процесу дозволяє обґрунтовано зіставляти режими роботи «ручний пошук», «чистий LLM» і «LLM + Pinecone (RAG)» та демонструвати виграти у швидкодії й точності [7, с. 152].

Висновки

Поєднання у ході дослідження розроблено підхід до інтеграції систем штучного інтелекту із семантичною векторною базою даних, що забезпечує високоточне семантичне зіставлення запитів із контентом і створює передумови для розвитку інтелектуальних аналітичних платформ нового покоління. Така архітектура забезпечує логічну послідовність, відтворюваність та прозорість процесу аналітичної обробки інформації.

Експериментальні результати підтвердили доцільність інтеграції LLM типу ChatGPT із SVDB для підвищення ефективності процесів бізнес-аналізу, що проявилось у скороченні часових витрат, підвищенні точності ранжування (Precision@k, nDCG@k) і достовірності відповідей порівняно з ручним пошуком та ізольованими LLM-моделями. Поєднання векторного пошуку з діалоговим

міркуванням мінімізувало втрати контексту й підвищило пояснюваність результатів. Практична цінність підходу полягає у створенні масштабованої методології обробки знань, придатної для бізнес-аналітики, конкурентної розвідки та e-commerce. Подальші дослідження доцільно спрямувати на автоматизацію контролю достовірності, багатомовну індексацію та підвищення безпеки корпоративних даних, що формує основу для гібридних ШІ-систем нового покоління.

Список літератури

1. Візнюк І., Буглай Н., Куцак Л., Поліщук А., Киливник В. Використання штучного інтелекту в освіті. *Modern Information Technologies and Innovation Methodologies of Education in Professional Training: Methodology Theory, Experience, Problems*, 2021. № 59. С. 14–22.
2. Грицишин В., Габрусєва Н. Штучний інтелект сьогодні і завтра. *Фундаментальні та прикладні проблеми сучасних технологій: тези Міжнародної науково-технічної конференції*. Тернопіль, 2020. С. 247–248. URL: <https://cutt.ly/EwDG4ks9> (дата звернення: 07.10.2025)
3. Денєсяк О., Паламарчук Є. Комплексна система прокторінгу в інформаційних технологіях аналізу контексту в системах оцінювання знань. *Вісник Вінницького політехнічного інституту*, 2021. № 6. С. 93–99.
4. Дроздова В. В., Рудніцька К. В., Росквас І. А. Інноваційні технології викладання іноземних мов в умовах зростаючого впливу технологій штучного інтелекту на освітні процеси. *Академічні візії*, 2023. № 26. С. 1–16.
5. Дутчак А. «Редактор телебачення в мережі Інтернет: терміни та функції». *Збірник праць НДІ пресосознавства*, Львів, 2017. Вип. 7(25). С. 326-340.
6. Козубай І. В. Комплексний аналіз впливу штучного інтелекту на викладання та вивчення іноземних мов. *Актуальні питання гуманітарних наук*, 2023. № 1 (63). С. 210–214.
7. Лебошина, Н., Черкашина, Ю., Харатян, Е. Методичні перспективи імплементації ШІ у навчанні англійської мови військових фахівців: імплементація технології штучного інтелекту у систему підготовки військових спеціалістів. *Сучасні дослідження з іноземної філології*, 2024. № 2(26). С. 145–152.
8. Хорошайло О. С., Кочергіна С. С. Використання штучного інтелекту для підвищення якості викладання іноземних мов у вищому закладі освіти. *Науковий часопис Національного педагогічного університету імені М. П. Драгоманова. Серія 5. Педагогічні науки: реалії та перспективи*, 2023. № 93. С. 123-127. URL: <https://doi.org/10.31392/NPU-nc.series5.2023.93.25>. (дата звернення: 05.09.2025).
9. A. Kargin and T. Petrenko, "Feeling Artificial Intelligence for AI-Enabled Autonomous Systems," in *Proc. of 2022 IEEE Global Conf. on Artificial Intelligence and Internet of Things (GCAIoT) Alamein New City, Egypt, Dec.18, 2022*, pp.88-93.
10. Beltagy I., Lo K., Cohan A. Longformer: The Long-Document Transformer // *arXiv preprint*, 2020. arXiv:2004.05150.
11. Bessen, J. E. «AI and Jobs: The Role of Demand». NBER Working Paper, №. 24235. National Bureau of Economic Research, 2019. pp. 14-15.
12. Davenport T. H., Kalakota, R. The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2), 2019. pp. 94-98.
13. David Bunting. How to use GenAI for database query optimization and natural language analysis. URL: <https://www.chaossearch.io/blog/genai-database-query-natural-language> (accessed Nov. 20, 2024).

14. E. Guizzo, Types of Robots. Categories frequently used to classify robots. Robotsguide.com. URL: <https://robotsguide.com/learn/types-of-robots> (accessed Jun. 15, 2025).
15. Lewis P. S. et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. arXiv preprint, 2020. arXiv:2005.11401.
16. Open X-Embodiment: Robotic Learning Datasets and RT-X Models Open - Embodiment Collaboration, 2024. [Online]. URL: <https://robotics-transformer-x.github.io/> (дата звернення: 10.10.2025).
17. Sapaty P, Morozov A, Finkelstein R, et al.) A new concept of flexible organization for distributed robotized systems. Proceedings of the 12th International Symposium on Artificial Life and Robotics (AROB 12), Beppu, Oita, Japan, Jan, 2007. 25–27, 2007, 8 p
18. Topol, E. Deep Medicine: ow Artificial Intelligence Can Make Healthcare Human Again / E. Topol. New York: Basic Books, Mar, 2019. pp.75–120.
19. Tserkovnyi V. O., Lytvynenko I. V. Analiz platform Arduino, Raspberry Pi ta ESP8266 v zadachakh monitorynhu perevezen bahazhu. Analysis of Arduino, Raspberry Pi and ESP8266 platforms in the tasks of baggage transportation monitoring. X I International scientific and practical conference of young researchers and students «Current issues in modern technologies» (Tern., 7-8 December 2022), 2022. pp. 59.
20. Variant Effect Prediction in the Age of Machine Learning — Y. Bromberg, R. Prabakaran, A. Kabir, A. Shehu. Cold Spring Harbor Perspectives in Biology. 2024. DOI:10.1101
21. Використання систем бізнес-аналітики в стратегічному управлінні підприємством (Гончарук О., Ландяк Т., Мельник Л.) — «Review of Transport Economics and Management», №13(29), 2025. DOI: 10.15802/rtem2025/339100.

References

1. Vizniuk I., Buhlai N., Kutsak L., Polishchuk A., Kylyvnyk V. Vykorystannia shtuchnoho intelektu v osviti. Modern Information Technologies and Innovation Methodologies of Education in Professional Training: Methodology Theory, Experience, Problems, 2021. № 59. S. 14–22.
2. Hrytsyshyn V., Habruseva N. Shtuchnyi intelekt sohodni i zavtra. Fundamentalni ta prykladni problemy suchasnykh tekhnolohii: tezy Mizhnarodnoi naukovo-tekhnichnoi konferentsii. Ternopil, 2020. S. 247– 248. URL: <https://cutt.ly/EwDG4ks9> (data zvernennia: 07.10.2025)
3. Denesiak O., Palamarchuk Ye. Kompleksna systema proktorinhu v informatsiinykh tekhnolohiiakh analizu kontekstu v systemakh otsiniuvannia znan. Visnyk Vinnytskoho politekhnichnoho instytutu, 2021. № 6. S. 93–99.
4. Drozdova V. V., Rudnitska K. V., Roskvas I. A. Innovatsiini tekhnolohii vykladannia inozemnykh mov v umovakh zrostaiuchoho vplyvu tekhnolohii shtuchnoho intelektu na osviti protsesy. Akademichni vizii, 2023. № 26. С. 1–16.
5. Dutchak A. «Redaktor telebachennia v merezhi Internet: terminy ta funktsii». Zbirnyk prats NDI presoznavstva, Lviv, 2017. Vyp. 7(25). S. 326-340.
6. Kozubai I. V. Kompleksnyi analiz vplyvu shtuchnoho intelektu na vykladannia ta vyvchennia inozemnykh mov. Aktualni pytannia humanitarnykh nauk, 2023. № 1 (63). S. 210–214.
7. Lieboshyna, N., Cherkashyna, Yu., Kharatian, E. Metodychni perspektyvy implementatsii Shl u navchanni anhliskoi movy viiskovykh fakhivtsiv: implementatsiia

tekhnologii shtuchnoho intelektu u systemu pidhotovky viiskovykh spetsialistiv. Suchasni doslidzhennia z inozemnoi filolohii, 2024. № 2(26). S. 145–152.

8. Khoroshailo O. S., Kocherhina S. S. Vykorystannia shtuchnoho intelektu dlia pidvyshchennia yakosti vykladannia inozemnykh mov u vyshchomu zakladi osvity. Naukovyi chasopys Natsionalnoho pedahohichnoho universytetu imeni M. P. Drahomanova. Serii 5. Pedahohichni nauky: realii ta perspektyvy, 2023. № 93. S. 123-127. URL: <https://doi.org/10.31392/NPU-nc.series5.2023.93.25>. (data zvernennia: 05.09.2025).

9. A. Kargin and T. Petrenko, "Feeling Artificial Intelligence for AI-Enabled Autonomous Systems," in Proc. of 2022 IEEE Global Conf. on Artificial Intelligence and Internet of Things (GCAIoT) Alamein New City, Egypt, Dec.18, 2022, pp.88-93.

10. Beltagy I., Lo K., Cohan A. Longformer: The Long-Document Transformer // arXiv preprint, 2020. arXiv:2004.05150.

11. Bessen, J. E. «AI and Jobs: The Role of Demand». NBER Working Paper, №. 24235. National Bureau of Economic Research, 2019. pp. 14-15.

12. Davenport T. H., Kalakota, R. The potential for artificial intelligence in healthcare. Future Healthcare Journal, 6(2), 2019. pp. 94-98.

13. David Bunting. How to use GenAI for database query optimization and natural language analysis. URL: <https://www.chaossearch.io/blog/genai-database-query-natural-language> (accessed Nov. 20, 2024).

14. E. Guizzo, Types of Robots. Categories frequently used to classify robots. Robotsguide.com. URL: <https://robotsguide.com/learn/types-of-robots> (accessed Jun. 15, 2025).

15. Lewis P. S. et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. arXiv preprint, 2020. arXiv:2005.11401.

16. Open X-Embodiment: Robotic Learning Datasets and RT-X Models Open X-Embodiment Collaboration, 2024. [Online]. URL: <https://robotics-transformer-x.github.io/> (data zvernennia: 10.10.2025).

17. Sapaty P, Morozov A, Finkelstein R, et al.) A new concept of flexible organization for distributed robotized systems. Proceedings of the 12th International Symposium on Artificial Life and Robotics (AROB 12), Beppu, Oita, Japan, Jan, 2007. 25–27, 2007, 8 p

18. Topol, E. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again / E. Topol. New York: Basic Books, Mar, 2019. pp.75–120.

19. Tserkovnyi V. O., Lytvynenko I. V. Analiz platform Arduino, Raspberry Pi ta ESP8266 v zadachakh monitorynhu perevezen bahazhu. Analysis of Arduino, Raspberry Pi and ESP8266 platforms in the tasks of baggage transportation monitoring. X I International scientific and practical conference of young researchers and students «Current issues in modern technologies» (Tern., 7-8 December 2022), 2022. pp. 59.

20. Variant Effect Prediction in the Age of Machine Learning — Y. Bromberg, R. Prabakaran, A. Kabir, A. Shehu. Cold Spring Harbor Perspectives in Biology. 2024. DOI:10.1101

21. Vykorystannia system biznes-analityky v stratehichnomu upravlinni pidpriemstvom (Honcharuk O., Landiak T., Melnyk L.) — «Review of Transport Economics and Management», №13(29), 2025. DOI: 10.15802/rtem2025/339100.

Надійшла до редакції 15.12.2025, розглянута на редколегії 11.02.2026

An approach to extending the functionality of artificial intelligence in the form of chatgpt using pinecone technology

The article addresses the problem of diversity of web interfaces, which affects the efficiency of business analysis, particularly in terms of time costs and the search for relevant information. The purpose of the study is to develop a theoretical and methodological framework for extending the functionality of artificial intelligence systems through the use of vector semantic database technology. The proposed approach is aimed at optimizing the processes of data search, structuring, and interpretation in business analysis, thereby increasing the productivity of business analysts and the quality of the results obtained.

The object of the research is business analysis processes using artificial intelligence systems, while the subject is the methodological and software support for their application based on vector semantic database technology. The core idea is to create a set of methods and software solutions that ensure effective human-machine interaction and reduce time costs in performing analytical operations.

In the course of the study, a critical review of the current state of the use of language models in business analytics was conducted, methods of human-machine interaction were improved, and an experimental prototype was developed to verify the effectiveness of the proposed solutions. The experimental results confirmed an increase in the efficiency of solving applied business analysis tasks using language models enhanced by semantic search and content matching technologies.

The obtained results can be applied in related fields of digital analytics, big data processing, e-commerce, and information flow management on the Internet. This increases the universality and practical value of the proposed approach, making it a promising direction for further development of intelligent analytical systems, their integration with semantic databases, and the enhancement of their adaptability to the dynamic web environment.

Keywords: AI; semantic search; artificial intelligence; business analysis; semantic vector databases; human-machine interaction; data optimization.

Відомості про авторів:

Феоктистова Олена Ігорівна – к.т.н., доцент, доцент кафедри інженерії програмного забезпечення, Національний аерокосмічний університет "Харківський авіаційний інститут". Електронна пошта: o.feoktystova@khai.edu, ORCID: <https://orcid.org/0000-0002-3051-0488>

Альошин Ярослав Михайлович – аспірант кафедри інженерії програмного забезпечення, Національний аерокосмічний університет "Харківський авіаційний інститут". Електронна пошта: y.aloshyn@khai.edu, ORCID: <https://orcid.org/0009-0006-0293-217X>

About the Authors:

Feoktystova Olena – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Software Engineering, National Aerospace University "Kharkiv Aviation Institute". Email: o.feoktystova@khai.edu, ORCID: <https://orcid.org/0000-0002-3051-0488>

Aloshyn Yaroslav – PhD student of the Department of Software Engineering, National Aerospace University "Kharkiv Aviation Institute". Email: y.aloshyn@khai.edu, ORCID: <https://orcid.org/0009-0006-0293-217X>