Р Е Ц Е Н З І Я

Міжнародний конкурс студентських робіт зі штучного інтелекту 2026

ІІ тур
Перший етап

на наукову роботу

назва Розробка і дослідження програмних інструментів для автоматизованої обробки текстових даних

шифр RAG-Lab

представлену на Міжнародний конкурс студентських робіт зі штучного інтелекту

№ з/п	Характеристики та критерії оцінки рукопису наукової роботи	Рейтингова оцінка. Максимальна кількість балів (за 120-бальною шкалою)	Бали
1	Актуальність проблеми	10	8,5
2	Новизна та оригінальність ідей	15	6,5
3	Використані методи дослідження	15	10,0
4	Теоретичні наукові результати	10	3,8
5	Практична направленість результатів	20	15,5
6	Рівень використання наукової літератури та інших джерел інформації	5	2,8
7	Ступінь самостійності роботи	10	8,0
8	Якість оформлення	5	4,3
9	Наукові публікації	10	3,3
10	Недоліки роботи (пояснення зниження максимальних балів у пунктах 1-9):		

Наукова новизна має переважно прикладний та експериментальний характер. Авторський внесок у порівняльне оцінювання конфігурацій локальної RAG-системи доцільно чіткіше відокремити від функціональності використаних бібліотек і відомих алгоритмів.

Методика дослідження потребує повнішого опису параметрів пошуку, фрагментації документів, квантування моделей і режимів виконання. Не наведено довірчих інтервалів та достатнього аналізу повторюваності вимірювань. Висновки про переваги конфігурацій слід обмежити дослідженими даними й апаратною платформою.

Оцінювання генерації на 50 прикладах дозволяє здійснити попереднє зіставлення моделей, проте недостатнє для стійкого ранжування близьких результатів. Використання різних моделей-суддів потребує перевірки узгодженості їхніх оцінок та вибіркової експертної валідації.

Теоретична частина переважно узагальнює відомі підходи. Твердження про гарантоване усунення фактологічних помилок завдяки RAG потребують уточнення, оскільки відповідну гарантію в роботі не доведено.

Огляд джерел доцільно доповнити первинними науковими публікаціями щодо гібридного пошуку та оцінювання RAG-систем. Наявні окремі мовні й редакційні неточності. Відомості про власні наукові публікації за результатами роботи не наведено.

Практичні результати бажано доповнити демонстрацією завершеного користувацького сценарію та перевіркою на незалежному корпусі документів.

Автор виконав великий обсяг практичної роботи: побудував працездатну локальну систему та провів послідовне порівняння багатьох моделей за якістю, швидкістю й ресурсами. Водночас: робота являє собою радше якісний технічний огляд-порівняння, ніж наукове дослідження: вибір компонентів обґрунтовано, однак нових підходів не запропоновано; вимірювання виконано на одному ноутбуці без повторів, тож частина відмінностей між моделями може пояснюватися випадковими чинниками; висновки варто обмежити цією конфігурацією; для оцінки відповідей використано мовні моделі як арбітрів на невеликій вибірці; бажано хоча б частково перевірити ці оцінки вручну; огляд спирається переважно на сайти, документацію та

10	форуми; посилань на рецензовані наукові праці мало; обіцяну інтеграцію з Telegram-ботом і роботу з реальними документами підприємства не показано; текст містить мовні огріхи. Проблема актуальна, проте задачі наясть суто інженерний характер. Новизна відсутня. Відбувається порівняння бібліотек і моделей. Наявність низки суттєвих недоліків, таких як мала вибірка оцінювання, методологічно некоректні експерименти. Робота не містить формалізованих моделей, математичних постановок задач чи теоретичних узагальнень. Система зібрана як інженерна зв'язка готових сервісів. Немає підтверджень практичної цінності. Фактично відсутні посилання на матеріали наукових публікацій. Робота виконана самостійно, проте складається з стандартних прикладів. Тема побудови RAG-системи є досить актуальною, але варто чіткіше визначити цільову предметну область і конкретні сценарії використання. Запропоноване рішення є корисним прикладом застосування RAG, однак рівень наукової новизни потребує чіткішого обґрунтування. Методологію проведення експериментів варто деталізувати, описати вибірку документів, спосіб розбиття на фрагменти, моделі вкладень, параметри векторного пошуку, формування контексту та процедуру генерації відповіді. Робота містить практичні та корисні результати, водночас теоретичне узагальнення є обмеженим. Доцільно формальніше описати критерії якості пошуку, вплив розміру фрагментів і глибини контексту на точність відповідей, а також межі застосування. RAG-система має високий потенціал для роботи з локальними базами знань і документами. Водночас для реального впровадження потрібно оцінити якість відповідей на різних типах запитів, частоту галюцинацій, поведінку за відсутності релевантного контексту Перелік джерел є достатнім просто для загального огляду та практично не містить наукоємних ресурсів. Його варто розширити новітніми публікаціями щодо RAG, chunking-стратегій, гібридного пошуку, reranking та оцінювання якості відповідей LLM. Робота демонструє самостійну реалізацію системи. Водночас слід чіткіше розмежувати власний внесок автора й використані готові LLM, бібліотеки, моделі вкладень, векторні БД та інструменти. Робота має структурований виклад, але для кращого сприйняття результатів доцільно додати більше графіків, таблиць. Доцільно детальніше описати архітектуру, технологічний стек, залежності та умови відтворення експериментів. Для посилення практичної цінності роботи варто навести кількісні метрики: точність, повноту, F1-міру, якість відповідей, частоту галюцинацій, час відповіді та порівняння з базовими RAG-рішеннями.		
10.1	Наявність розробленого програмного забезпечення, у т.ч. веб-сервісу, мобільного застосунку	10	7,3
10.2	Рівень отриманих практичних результатів наукової роботи	10	6,5
Сума балів			76,3

Загальний висновок не рекомендується для захисту на ПНПК

(рекомендується, не рекомендується для захисту на науково-практичній конференції)

Рецензент *члени Конкурсної комісії напряму «Обробка природної мови»*

«06» листопада 2026 року