

СТВОРЕННЯ ІНТЕЛЕКТУАЛЬНОГО ВЕБ-АСИСТЕНТА ДЛЯ АНАЛІЗУ ДАНИХ З КУЛЬОВОЇ СТРІЛЬБИ НА ОСНОВІ LLM

Вінницький національний технічний університет

Анотація

Розглянуто розробку інтелектуального веб-додатка для збору, агрегації та аналізу актуальної інформації з тематики кульової стрільби із використанням великих мовних моделей (LLM). Запропонована система поєднує технології генерації з доповненою вибіркою (RAG), глибокого веб-парсингу офіційних джерел та глобального веб-пошуку для формування комплексних звітів у режимі реального часу. Реалізовано інтеграцію моделі Gemini 2.5 Flash через API та розроблено модуль інтерактивного чату зі збереженням контексту сесії на базі фреймворку Streamlit. Описано архітектуру гібридного пошуку, основні функціональні модулі та результати тестування програмного забезпечення.

Ключові слова: великі мовні моделі, веб-парсинг, кульова стрільба, генерація з доповненою вибіркою (RAG), штучний інтелект.

Abstract

The development of an intelligent web application for the collection, aggregation, and analysis of up-to-date information on target shooting using large language models (LLMs) is considered. The proposed system combines Retrieval-Augmented Generation (RAG) technology, deep web scraping of official sources, and global web search to generate comprehensive reports in real time. The integration of the Gemini 2.5 Flash model via API has been implemented, and an interactive chat module with session context retention has been developed based on the Streamlit framework. The hybrid search architecture, main functional modules, and software testing results are described.

Keywords: large language models, web scraping, target shooting, retrieval-augmented generation (RAG), artificial intelligence.

Вступ

Інформаційні технології на базі великих мовних моделей (LLM) відкривають нові можливості для створення вузькоспеціалізованих аналітичних систем. На відміну від універсальних чат-ботів, такі системи здатні працювати з глибоким контекстом конкретної предметної області [1]. Розробка подібних рішень має не лише прикладне, а й вагоме теоретичне значення, оскільки слугує експериментальним підґрунтям для вдосконалення LLM із застосуванням принципів асоціативного образного мислення людини [2, 3].

Спортивна аналітика у сфері кульової стрільби вимагає надзвичайної точності даних: відстеження результатів турнірів, змін у технічному регламенті та статистики світових рекордів. Оскільки протоколи змагань (зокрема під егідою ISSF [4]) оновлюються безперервно, базові мовні моделі часто генерують застарілу інформацію або допускають фактологічні помилки («галюцинації»). Вирішенням цієї проблеми є інтеграція LLM із методами динамічного веб-парсингу та архітектурою RAG (Retrieval-Augmented Generation) [5].

Метою роботи є створення інтелектуальної веб-системи, яка автоматизує збір, комплексний аналіз та інтерактивний пошук актуальної інформації зі спортивної кульової стрільби на рівні державних та світових досягнень.

Результати дослідження

Програмний комплекс реалізовано мовою Python із застосуванням фреймворку Streamlit для побудови інтерактивного користувацького інтерфейсу [6]. Основою архітектури системи є RAG-підхід, який підвищує достовірність відповідей ШІ шляхом надання моделі жорсткого фактологічного контексту із зовнішніх джерел [5]. Для обробки текстів та формування звітів залучено модель Gemini 2.5 Flash, яка оптимізована для швидкого аналізу великих масивів неструктурованих даних. Збір інфор-

мації здійснюється за допомогою розробленого алгоритму гібридного пошуку. Першим рівнем є глибокий веб-парсинг (Deep Scraping) офіційних джерел, зокрема сайту Федерації стрільби України та новинного порталу ISSF [4]. Другим рівнем є глобальний веб-пошук. Для розширення інформаційної бази та забезпечення доступу до світових новин у реальному часі в систему інтегровано програмну бібліотеку пошукової системи DuckDuckGo [7]. Використання DuckDuckGo дозволяє виконувати автоматизовані пошукові запити без потреби у складних API-ключах та агрегувати найрелевантніші результати зі спортивних медіа, компенсуючи можливі затримки публікацій на офіційних сайтах [7]. Особливу увагу під час розробки було приділено подоланню «часових галюцинацій» моделі. Шляхом інженерії підказок (Prompt Engineering) [8] реалізовано логічний контроль дат, що дозволило системі безпомилково розрізняти минулі, поточні та майбутні спортивні події відносно системного часу. Приклад інтерфейсу аналітичної веб-системи та приклад генерації базового звіту наведено на рис. 1.

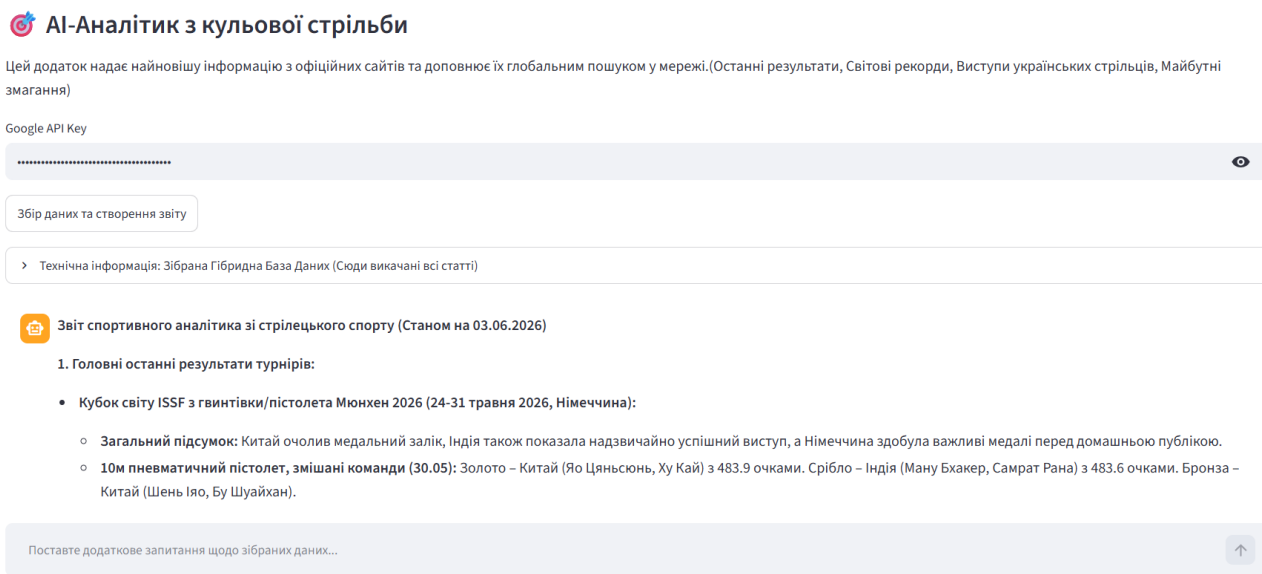


Рисунок 1 – Інтерфейс аналітичної веб-системи та приклад генерації звіту

Зібрані через парсинг та DuckDuckGo тексти об'єднуються у єдиний інформаційний пул. Після цього система автоматично генерує структурований звіт про останні змагання, рекорди та досягнення спортсменів.

Для забезпечення розширеної взаємодії в системі реалізовано модуль інтерактивного чату. Завдяки механізму управління станом сесії (Session State), програма зберігає контекст попередніх повідомлень та всю завантажену базу статей. Для максимальної точності відповідей у модуль чату інтегровано алгоритм динамічного пошуку (Dynamic Retrieval) [4, 7]. При кожному новому запитанні користувача система не лише звертається до зібраної бази, а й здійснює миттєвий цільовий пошук в Інтернеті. Це дозволяє користувачу ставити системі уточнювальні запитання щодо конкретних очок стрільців чи деталей фіналів, отримуючи миттєві відповіді на основі найсвіжіших фактів без ризику генерації відмовок з боку ШІ. Фрагмент роботи інтерактивного чату наведено на рис. 2.

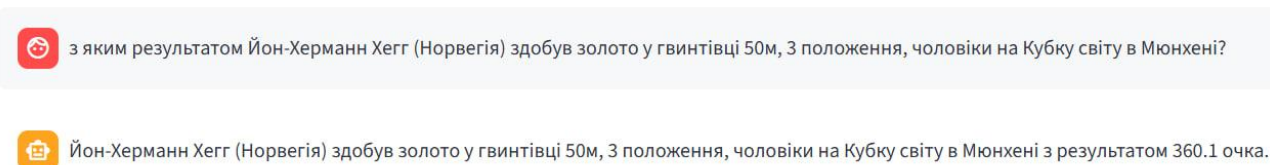


Рисунок 2 – Робота інтерактивного чату із застосуванням динамічного глобального веб-пошуку

Висновки

У результаті роботи було розроблено інтелектуальну аналітичну веб-систему з тематики кульової стрільби на основі великої мовної моделі Gemini. Реалізована архітектура забезпечує збір, глибокий парсинг та аналіз актуальної інформації з офіційних спортивних ресурсів у режимі реального часу. Використання розширеної технології Retrieval-Augmented Generation, що поєднує пряме зчитування веб-сторінок із динамічним пошуковим індексуванням, дозволило нівелювати основні недоліки класичних LLM. Впроваджені алгоритми логічного контролю часу та динамічного пошуку в чаті, реалізовані за допомогою методів інженерії підказок [8], забезпечили високу точність та фактологічну достовірність згенерованих звітів, перетворивши систему на актуальний та перспективний інструмент для спортивних аналітиків.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Brown T. Language Models are Few-Shot Learners / T. Brown, B. Mann, N. Ryder et al. // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 1877-1901.
2. Омельченко, В.; Бісікало, О. Удосконалення великих мовних моделей формальними засобами оцінки сенсу текстової інформації. ВІТКІП ВНТУ. Факультет інтелектуальних інформаційних технологій та автоматизації, Ukraine, mar. 2026. Available at: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2026/index> . (дата звернення: 03.06.2026).
3. Бісікало О. В. Формальні методи образного аналізу та синтезу природно-мовних конструкцій : монографія. — Вінниця : ВНТУ, 2013. – 240 с.
4. International Shooting Sport Federation (ISSF) Official Website [Електронний ресурс]. – Режим доступу: <https://www.issf-sports.org/> (дата звернення: 03.06.2026).
5. Lewis P. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks / P. Lewis, E. Perez, A. Piktus et al. // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 9459–9474.
6. Bispo N. Building Web Apps with Streamlit: A Guide for Data Enthusiasts / N. Bispo // Medium. – 2023 [Електронний ресурс]. – Режим доступу: <https://medium.com/django-unleashed/building-web-apps-with-streamlit-a-guide-for-data-enthusiasts-e805452a0ab> (дата звернення: 03.06.2026).
7. DuckDuckGo Search Documentation [Електронний ресурс]. – Режим доступу: <https://duckduckgo.com/> (дата звернення: 03.06.2026).
8. Liu P. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing / P. Liu, W. Yuan, J. Fu et al. // ACM Computing Surveys. – 2023. – Vol. 55, No. 9. – P. 1–35.

Бісікало Олег Володимирович – д-р техн. наук, професор, завідувач кафедри автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, м. Вінниця, email: obisikalo@vntu.edu.ua

Собчук Ольга Олександрівна – студентка групи ІІСТ-24б, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, email: o.sobchuk878@gmail.com

Bisikalo Oleh V. – Dr. Sc. (Eng.), Professor, Head of the Department of Automation and Intelligent Information Technologies, Vinnytsia National Technical University, Vinnytsia, email: obisikalo@vntu.edu.ua

Sobchuk Olha O. – student of group IIIST-24b, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, email: o.sobchuk878@gmail.com