

## СИСТЕМИ З ВИКОРИСТАННЯМ LLM ЯК МОТИВАТОР ДЛЯ НАВЧАННЯ

Небаба М.Ю., Рябова Н.В.

e-mail: maksym.nebaba@nure.ua, nataliya.ryabova@nure.ua

Харківський національний університет радіоелектроніки, каф. ШІ,  
м. Харків, Україна

This work investigates the integration of artificial intelligence with multimedia animation in virtual streaming platforms. The system architecture is based on a large language model (LLM). Real-time text generation and adaptive linguistic modulation are achieved through advanced algorithmic techniques. Synchronized voice and animated avatar interactions are facilitated by efficient API-based connectivity. Reviewed Neuro-sama system, mentioned opensource platforms and APIs, and how they encourage learning and experimenting with machine learning and data science.

Революційний розвиток генеративного штучного інтелекту (Generative Artificial Intelligence, GenAI), який розпочався декілька років тому, вже досяг того рівня коли гаслом стало висловлення «generation from anything to anything». Тобто з'являється все більше проектів, в яких генеруються задані типи об'єктів із різноманітних джерел. В той же час розвиток GenAI супроводжується появою і постійним поповненням великих мовних моделей (Large Language Model, LLM), які перетворилися на окремих напрямках досліджень та прикладних розробок [1]. Незважаючи на достатньо велику кількість вже розроблених LLM, які знаходяться у вільному доступі, їх кількість постійно зростає, а якість – поліпшується. Найбільш ефективним підходом до створення LLM став підхід на основі найбільш передових архітектур глибоких нейронних мереж – трансформерів, основною парадигмою яких є «Attention is all you need» [2]. Однак, одним з найбільш актуальних та розповсюджених застосувань GenAI та LLM сьогодні є розробка різного роду прикладних систем із використанням великих мовних моделей.

Дана робота спрямована на дослідження та використання технологій GenAI та найбільш перспективних моделей LLM для побудови прикладної інтелектуальної системи, спроможної до генерації контексту, відео та ведення діалогу у реальному часі із використанням аватара (інтелектуального агента). В якості предметної області розглядається розважальна індустрія, а саме, стрімінгові платформи для розваг глядачів та гравців. Необхідно розробити ШІ-стрімера із власним аватаром, тобто нейронну мережу, яка може повністю автономно вести стрім, утримуючи увагу глядачів. Основною вимогою до системи є спроможність повністю автономно і без перерви вести трансляцію на обраній стрімінг-платформі, розважаючи аудиторію.

У сучасних технологічних системах, що поєднують штучний інтелект із мультимедійною анімацією, Neuro-sama виступає як яскравий приклад системи з використанням ШІ для стрімінгу. Архітектура даної системи побудована на основі великої мовної моделі, яка дозволяє генерувати текст в режимі реального часу. Завдяки використанню трансформерів, система здатна адаптуватися до запитів аудиторії, що дозволяє підтримувати живу та природну комунікацію у розмові. Технічна реалізація включає розробку алгоритмів обробки природної мови (Natural Language Processing, NLP), що дозволяють генерувати відповіді на основі контексту розмови, а також системи синхронізації голосу із анімацією аватара, завдяки чому голосовий супровід відображається відповідними рухами обличчя та мімікою.

Інтеграція з платформою Twitch реалізована через використання API, що забезпечує оперативний обмін даними між серверною частиною та інтерфейсом системи. Це дозволяє швидко реагувати на повідомлення аудиторії та координувати взаємодію між різними компонентами, включаючи обробку аудіо та відео-потоків. Реалізовані алгоритми з низькою затримкою забезпечують синхронність роботи між голосовим модулем та анімованим персонажем, що є критично важливим для створення переконливої і природньої взаємодії.

З технічної сторони, Neuro-sama працює за принципами мультимодальної інтеграції, де різні типи даних, такі як текст, аудіо та відео, об'єднані в єдину систему. Такий підхід дозволяє не лише підтримувати діалог, але й забезпечує візуальне та емоційне зображення персонажу, що підсилює ефект присутності віртуального стрімера. В системі використовуються алгоритми для фільтрації та модерації контенту. Ці алгоритми, засновані на багаторівневих підходах до цензурування, де комбінація попередньо навчених моделей та алгоритмічної перевірки допомагає уникнути небажаних висловлювань або непередбачених ситуацій.

Теоретична концепція Neuro-sama ґрунтується на розвитку трансформерних архітектур, здатних обробляти великі обсяги даних і адаптуватися до нових умов в режимі реального часу. Це дозволяє системі аналізувати численні мовні патерни та використовувати їх для генерації відповідей, що є основою для реалізації інтерактивних голосових помічників та віртуальних аніматорів.

Не менш важливим аспектом є використання доступних API, таких як OpenAI, що дає змогу розробникам швидко створювати прототипи та експериментувати з новими формами контенту. Окрім OpenAI можна відмітити платформу Hugging Face, яка надає великий спектр моделей для вирішення задач комп'ютерного зору, обробки аудіо, NLP та різноманітних баз даних. Такі платформи відкривають можливості для розробників-початківців, та дозволяють створювати власних віртуальних ведучих або стрімерів із застосуванням сучасних алгоритмів NLP.

Внаслідок цього відбувається розширення практичного досвіду у галузі, а користувачі, що спершу залучені розважальним форматом, згодом переходять до більш детального розуміння принципів роботи нейронних мереж.

Важливо відмітити, що відкриті проекти типу Neuro-sama є не тільки розважальними. Інтерактивність та доступність технології сприяють тому, що аудиторія не лише спостерігає за роботою системи, але й активно досліджує її внутрішні механізми, що розширює розуміння можливостей і обмежень сучасних технологій. Таким чином, приклад Neuro-sama демонструє, як розважальний проект може стати платформою для навчання та інновацій, сприяючи появі нових розробників та дослідників у галузі штучного інтелекту.

У даній роботі передбачається розробка власного ШІ-стрімера з основним функціоналом, подібним Neuro-sama, але адаптованого до обраної предметної галузі. Перспективний розвиток подібних інтегрованих систем може призвести до появи ще більш складних та адаптивних платформ, які об'єднують у собі високопродуктивні алгоритми обробки даних, синхронізовану мультимедійну взаємодію та глибокі механізми модерації контенту. Це створить умови для формування нового покоління технологій, інтерактивні системи стануть невід'ємною частиною як цифрової культури, так і процесів освітнього та технологічного розвитку.

#### Список використаних джерел:

1. Bhawna Singh. Building Applications with Large Language Models. Techniques, Implementations, and Applications. APRESS, 2024.
2. Rashka Sebastian. Building a Large Language Model (From Scratch). Manning, 2025.
3. Contributors to Wikimedia projects. Neuro-sama – wikipedia. Wikipedia, the free encyclopedia. URL: <https://en.wikipedia.org/wiki/Neuro-sama> (date of access: 04.03.2025).
4. Streamer AI (Like Neuro-Sama). Hugging Face Forums. URL: <https://discuss.huggingface.co/t/streamer-ai-like-neuro-sama/33836> (date of access: 04.03.2025).
5. This virtual twitch streamer is controlled entirely by AI. VICE. URL: <https://www.vice.com/en/article/this-virtual-twitch-streamer-is-controlled-entirely-by-ai/> (date of access: 04.03.2025).