

ГЕНЕРАЦІЯ ТЕСТОВИХ ЗАВДАНЬ З ТЕКСТОВИХ НАВЧАЛЬНИХ МАТЕРІАЛІВ

Гуржи Т.С., Гриньова О.Є.

e-mail: tetiana.hurzhy@nure.ua, olena.hrynova@nure.ua

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

The development of an automated test generation system using large language models (LLM) such as GPT and Perplexity (Sonar) offers significant advancements in education. This innovative system automates the creation of test questions from various educational materials, generates realistic distractors, and classifies questions by difficulty level. By enhancing assessment objectivity and reducing educator workload, it creates more effective tests while ensuring comprehensive evaluation of students' knowledge and critical thinking skills.

Штучний інтелект (ШІ) все більше впливає на освітній процес, відкриваючи нові можливості для автоматизації різних завдань [1]. Одним із таких завдань є створення тестових завдань для перевірки знань студентів, що, зазвичай, займає значну кількість часу у викладачів. З появою великих мовних моделей (LLM) стало можливим автоматизувати цей процес, підвищуючи його ефективність та знижуючи навантаження на педагогів.

Однією з головних проблем у тестуванні знань студентів є те, що вони можуть вгадати правильну відповідь, не маючи глибокого розуміння матеріалу. Це особливо актуально для тестів з вибором відповіді, де студенти часто використовують метод виключення або інтуїцію, а не реальні знання. У такому випадку тест не виконує свою основну функцію – об'єктивно оцінювати рівень засвоєного матеріалу.

Крім того, поширення LLM, наприклад ChatGPT, дає студентам можливість отримувати відповіді на питання без необхідності самостійного аналізу матеріалу. Це створює новий виклик для освітнього процесу, як перевірити, чи студент дійсно розуміє тему, а не просто використовує ШІ.

Запропонована інтелектуальна система автоматично генерує тестові завдання на основі завантажених навчальних матеріалів, створюючи нові питання, формує варіанти відповідей та адаптує рівень складності до особливостей навчального контенту. Вона використовує передові моделі штучного інтелекту, які аналізують текст, генерують запитання та пропонують як правильні, так і неправдиві, але дуже реалістичні відповіді. Важливою особливістю системи є те, що кожна відповідь має свою вагу - числове значення від 0 до 1, яке вказує на ступінь правильності відповіді. Це дозволяє більш точно оцінювати знання студентів, враховуючи не лише абсолютно правильні відповіді, але й частково правильні.

Розроблена система класифікує питання за трьома рівнями складності (легкий, середній та складний), що дозволяє створювати збалансовані тести, які відповідають різним рівням підготовки студентів. Такий підхід не лише прискорює створення тестів та підвищує їхню якість, але й сприяє об'єктивності оцінювання, знижує ризик суб'єктивних помилок та ускладнює механічне вгадування відповідей. Це робить процес контролю знань більш ефективним та точним, забезпечуючи комплексну оцінку навичок студентів.

Одним із ключових методів покращення тестування є використання алгоритмічно згенерованих дистракторів — варіантів відповідей, які є неправдивими, але виглядають правдоподібно. Такі дистрактори повинні бути неочевидними, що змушує студента ретельно аналізувати питання, а не просто спиратися на інтуїцію чи шаблонні знання.

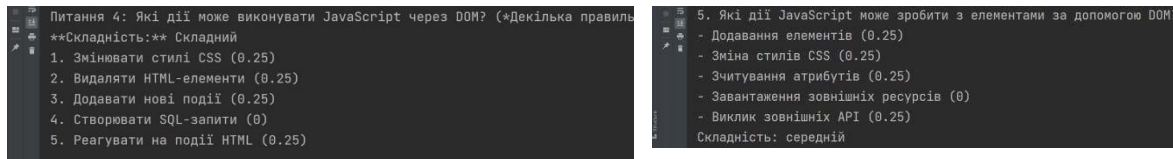
Генеративні моделі можуть ефективно створювати такі дистрактори завдяки механізму "галюцинацій", коли нейромережа генерує відповіді, що на перший погляд здаються коректними, але містять логічні або змістовні помилки. Це дозволяє оцінити, наскільки глибоко студент розуміє матеріал, а не просто розпізнає знайомі терміни.

Процес роботи системи передбачає завантаження навчальних матеріалів у зручному для викладача форматі (текст, презентація, файли формату PDF та Word). Після опрацювання матеріалу система генерує тестові завдання, які викладач може переглянути, відредагувати та адаптувати до своїх вимог. Інтуїтивно зрозумілий інтерфейс дозволяє легко користуватися системою навіть тим, хто немає досвіду роботи з цифровими інструментами.

Практична значимість системи полягає у її здатності автоматизувати створення тестів, зменшуючи навантаження на викладачів. Це особливо важливо в умовах дистанційного навчання, коли необхідність швидкої та якісної перевірки поточних знань є ключовим фактором. Використання дистракторів та механізму "галюцинацій" дозволяє створювати більш реалістичні, але водночас складніші для вгадування тестові завдання, що змушує студентів ретельно аналізувати кожен варіант відповіді. У майбутньому система може бути інтегрована із сучасними системами управління навчанням (LMS), такими як Moodle або Google Classroom, що дозволить забезпечити її широке впровадження у навчальні заклади та ефективну взаємодію з уже існуючими інструментами освітнього процесу.

У ході дослідження було проаналізовано два варіанти мовних моделей – ChatGPT та Perplexity (Sonar) [2, 3], кожен з яких продемонстрував свої унікальні переваги та обмеження. Враховуючи отримані результати, перспективним напрямком подальшого розвитку системи може стати комбінування підходів обох моделей для досягнення оптимального балансу між якістю та ефективністю генерації тестів (рисунок 1). Така інтегрована технологія не лише сприятиме розвитку

персоналізованого навчання, але й забезпечить більш ефективне використання цифрових інструментів в освітньому процесі.



а) Perplexity

б) ChatGPT

Рисунок 1 – Генерація тесту різними моделями LLM де а) результат роботи Perplexity, б) результат роботи ChatGPT.

Крім того, було б доцільно розширити можливості системи, додавши функціонал для створення тестів не лише на основі текстових матеріалів, але й з аудіо- та відеоформатів. Це дозволить викладачам використовувати більш різноманітні джерела інформації для генерації тестових завдань, що зробить процес оцінювання знань більш всебічним та адаптованим до сучасних форматів навчальних матеріалів. Система може надати студентам можливість самостійно створювати тести для практики та перевірки власних знань. Це дозволить їм не лише відповідати на запитання, а й активно формувати навчальний процес, аналізуючи свої помилки та прогалини в розумінні матеріалу. Подібна функція може стати частиною інтегрованого навчального середовища, де студенти взаємодіють із контентом більш глибоко, а викладачі отримують додаткову інформацію про теми, які потребують доопрацювання. Такий підхід сприятиме розвитку самостійності студентів та допоможе зробити навчальний процес більш адаптивним і персоналізованим.

Система може розширити свої можливості, надаючи студентам додаткові інструменти. Одним із таких реалізованих інструментів є автоматично згенерований глосарій курсу з ключовими термінами та їх визначеннями, що допоможе студентам краще орієнтуватися в темі та ефективніше готуватися до іспитів. Проте важливо зазначити, що експертиза викладача залишається незамінною у навчальному процесі, і не можна повністю покладатися на штучний інтелект. Викладач відіграє ключову роль у перевірці та доповненні автоматично згенерованого контенту, забезпечуючи його точність та релевантність для курсу.

Список використаних джерел:

1. Жаркіх С. Є., Гриньова О. Є. // Архітектура голосового аналітика під час вивчення іноземної мови // 28-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті». Зб. матеріалів форуму. Т. 6.1 – Харків: ХНУРЕ. 2024. – С.16-18. <https://doi.org/10.30837/IYF.IIS.2024.016>
2. OpenAI. API-документація: Chat Completions. URL: <https://platform.openai.com/docs/api-reference/chat/completions>.
3. Perplexity AI. API-документація: Chat Completions. URL: <https://docs.perplexity.ai/api-reference/chat-completions>.