

УДК 004.8:371.26

ВИКОРИСТАННЯ МАТЕМАТИЧНИХ І ШІ-МОДЕЛЕЙ ДЛЯ ГЕНЕРАЦІЇ ТЕСТІВ (LLM, ЕВОЛЮЦІЙНІ АЛГОРИТМИ, RL)

Шкурко В.В., Поляков А.О.

Харківський національний університет радіоелектроніки, Харків

Сучасний розвиток технологій штучного інтелекту (ШІ) зумовив появу нової генерації освітніх систем, здатних не лише аналізувати навчальний контент, а й автоматично створювати тестові завдання з урахуванням рівня складності, когнітивних цілей та індивідуальних характеристик користувача, що сформувало нові можливості для побудови адаптивних інтелектуальних систем оцінювання із поєднанням математичних моделей і ШІ-моделей – зокрема, великих мовних моделей (LLM), еволюційних алгоритмів і методів навчання з підкріпленням (Reinforcement Learning, RL).

Математичне моделювання завжди було основою для побудови точних і надійних рішень у сфері штучного інтелекту [5]. Сьогодні, коли великі мовні моделі на кшталт ChatGPT навчилися не лише обробляти текст, а й мислити послідовно, стало очевидно: ШІ поступово опановує логіку, яка раніше вважалася виключно людською. У роботі К. Шпрейцер, О. Штразер, С. Цехетмаєр та К. Маасс [1] показано, що такі моделі можуть не просто відтворювати хід математичних міркувань, а й пояснювати його, створюючи зрозумілі та навчальні приклади. Подібну закономірність помітив і С. Оу [2], зазначивши, що моделі *o1-preview* та *GPT-4o* демонструють вражаючу здатність розпізнавати математичні структури й утримувати логіку рішення, хоча іноді все ще помиляються у дрібних деталях, де потрібне людське уточнення. Водночас, як наголошують Р. Безханчер, В. Демір, ефективне застосування ШІ-інструментів у навчанні передбачає не лише технічну реалізацію, а й розвиток техно-математичної грамотності та самоефективності користувачів [6].

Інтеграція подібних систем у навчальний процес допомагає студентам не просто користуватися технологіями, а розуміти їхню логіку, розвиваючи цифрову компетентність і критичне мислення під час взаємодії з алгоритмами. Українські дослідники Н. Круглова та О. Диховичний [10] підкреслюють, що використання ІКТ у створенні програм-шаблонів для автоматизованої генерації тестових завдань з вищої математики є важливим етапом на шляху до впровадження інтелектуальних рішень на базі штучного інтелекту.

Поєднання математичних моделей із сучасними LLM та RL створює нове покоління адаптивних тестових систем, розвиваючись на стику математики, штучного інтелекту та педагогіки, що формує нову можливість підлаштовуватися до потреби користувача.

Серед провідних напрямів розвитку варто виокремити онтологічно-семантичні, еволюційні та RL-орієнтовані моделі. Один із прикладів подають П. Шарпф, М. Шуботц, А. Шпіц, А. Грайнер-Петтер і Б. Гіпп [4], які розробили колаборативну систему, що формує екзаменаційні питання, спираючись на бази знань *Wikidata*. Вона використовує механізми семантичного пошуку, аби добирати найточніші формулювання. Подібний принцип можна застосовувати й у сучасних освітніх платформах, де LLM працюють разом із навчальними базами, формуючи завдання, які відповідають контексту теми й рівню складності, але водночас залишаються зрозумілими й логічними для студента.

М. Анджеєвський, Н. Дубицька, Дж. Подоляк, М. Коваль і Дж. Сілка [9] запропонували алгоритмічну схему, яка передбачає чотири етапи: аналіз навчального контенту, побудову завдання, оцінку якості питання та оптимізацію відповідей. Кожен етап підкріплюється RL-циклами зворотного зв'язку, що дають змогу системі навчатися на власних результатах і поступово підвищувати педагогічну точність та логічну узгодженість питань.

Наукова спільнота активно досліджує можливості автоматизації створення тестів, проте водночас усе більше уваги приділяється питанню

надійності, валідності та довіри до таких систем. К. Ф. Гарсія, А. К. С. Онг, М. Дж. Дж. Гумасінг і С. Р. В. Делос Рейес [3] підкреслюють, що сприйняття ШІ-інструментів студентами прямо залежить від рівня довіри до системи та її здатності пояснювати логіку своїх рішень. Це зумовлює потребу у прозорих механізмах перевірки, валідації й пояснюваності результатів. Й. Мохамату, А. Халідоу та Р. Т. Капен [5] звертають увагу на те, що успішність поєднання математичних моделей і ШІ-алгоритмів визначається правильним налаштуванням параметрів валідації та оптимізації. Такий підхід дає змогу створити самонавчальні системи, які автоматично аналізують статистичні показники тестів – дискримінативність, надійність, складність – і коригують завдання за допомогою еволюційних алгоритмів, забезпечуючи безперервне вдосконалення якості контенту.

Р. Беязханчер, Б. Демір розглядають ще один важливий аспект – AI self-efficacy користувачів, тобто впевненість у власній здатності взаємодіяти з інтелектуальними системами [6]. Підвищення цього рівня прямо впливає на ефективність сприйняття тестів, створених ШІ, і сприяє розвитку техно-математичної грамотності студентів. Водночас, як зазначають Р. Ворса, Т. Тамаклоу, освітні AI-платформи повинні враховувати когнітивні особливості користувачів та уникати надмірної автоматизації, яка знижує мотивацію до мислення [8]. Оптимальним є поєднання алгоритмічного підбору завдань і можливості для студента рефлексувати над процесом вирішення. Отже, якість і валідність системи генерації тестів визначаються не лише алгоритмічною точністю, а й педагогічною релевантністю та психологічною прийнятністю. Комплексна модель має враховувати когнітивні, технічні й етичні чинники, створюючи простір для гармонійної взаємодії людини й штучного інтелекту у сфері оцінювання знань.

Сучасні наукові дослідження демонструють широку варіативність підходів до автоматизованого створення тестових завдань, що базуються на поєднанні методів оброблення природної мови, математичного моделювання, баз знань і навчання з підкріпленням. Попри спільну мету – підвищення

ефективності формування навчального контенту – ці моделі відрізняються рівнем автономності, принципами оцінювання якості та масштабованістю. На основі аналізу робіт [4; 7; 9; 10] узагальнено ключові характеристики сучасних підходів (табл. 1).

Таблиця 1 – Порівняльна характеристика сучасних підходів до автоматизованої генерації тестів із використанням ШІ-моделей

№	Автори, рік	Метод / модель	Джерело даних	Особливості реалізації	Переваги	Обмеження
1	Scharpf et al., 2022	Колаборативна генерація з Wikidata	Семантичні бази знань	Формування екзаменаційних запитань через онтологічний пошук і структуровані дані	Контекстна точність, відтворюваність результатів	Залежність від актуальності бази знань
2	Mead & Zhou, 2024	LLM-генерація сертифікаційних питань	Текстові корпуси та стандартизовані екзамени	GPT-моделі створюють питання і проводять внутрішню валідацію варіантів відповідей	Висока якість мовного формулювання	Обмежена адаптація до різних предметних областей
3	Andrzejewski et al., 2025	Автоматизована генерація з RL-циклом	Освітні матеріали й метадані тестів	Система самонавчається, оптимізує параметри якості (релевантність, складність, валідність)	Автокорекція завдань, адаптивність	Високі обчислювальні витрати
4	Круглова, Диховичний, 2025	ІКТ-шаблони для створення завдань	Навчальні ресурси з вищої математики	Генерація тестів через параметричні шаблони у поєднанні з ML-алгоритмами	Простота інтеграції, можливість локальної модифікації	Обмежена семантична варіативність, потреба ручного контролю

Проведений огляд і систематизація наукових джерел засвідчують, що використання математичних і ШІ-моделей для генерації тестів формує новий етап розвитку освітніх технологій, де аналітична точність математичних методів поєднується з гнучкістю генеративних мовних моделей, які, у свою чергу, забезпечують динамічну адаптацію рівня складності завдань, можливість самонавчання системи й підвищення педагогічної релевантності контенту. Проблематикою їх повсюдного впровадження є якість і валідність. Подальші дослідження мають бути спрямовані на розроблення гібридних архітектур, які забезпечуватимуть адаптивність, етичність та науково обґрунтованість тестування – це у перспективі відкриває можливості для створення інтелектуальних систем оцінювання нового покоління.

Література:

1. Spreitzer C., Straser O., Zehetmeier S., Maaß K. Mathematical Modelling Abilities of Artificial Intelligence Tools: The Case of ChatGPT. *Education Sciences*. 2024. Vol. 14, No. 7. P. 698. DOI: <https://doi.org/10.3390/educsci14070698>
2. Oh S. Evaluating Mathematical Problem-Solving Abilities of Generative AI Models: Performance Analysis of o1-preview and gpt-4o Using the Korean College Scholastic Ability Test. *IEEE Access*. 2024. Vol. 13. P. 1227–1235. DOI: <https://doi.org/10.1109/ACCESS.2024.3523703>
3. Garcia K. F., Ong A. K. S., Gumasing M. J. J., Delos Reyes C. R. V. Engineering students' perceptions and actual use of AI-based math tools for solving mathematical problems. *Acta Psychologica*. 2025. Vol. 256. Article ID: 105004. ISSN 0001-6918. DOI: <https://doi.org/10.1016/j.actpsy.2025.105004>
4. Scharpf P., Schubotz M., Spitz A., Greiner-Petter A., Gipp B. Collaborative and AI-aided Exam Question Generation using Wikidata in Education. *CEUR Workshop Proceedings. Wikidata'22: Wikidata workshop at ISWC 2022*. 2022. Vol. 3262. ISSN 1613-0073. URL: <https://ceur-ws.org/Vol-3262/paper13.pdf>

5. Mohamadou Y., Halidou A., Kapen P. T. A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. *Applied Intelligence*. 2020. Vol. 50, No. 11. P. 3913–3925. DOI: <https://doi.org/10.1007/s10489-020-01770-9>
6. Beyazhancer R., Demir B. Enhancing techno-mathematical literacy and AI self-efficacy in engineering education through artificial intelligence applications. *Frontiers in Education*. 2025. Vol. 10. Article ID: 1695351. DOI: <https://doi.org/10.3389/feduc.2025.1695351>
7. Mead A. D., Zhou C. Evaluating the Quality of AI-Generated Items for a Certification Exam. *Journal of Applied Testing Technology*. 2024. URL: <https://jattjournal.net/index.php/atp/article/view/173204>
8. Vorsah R., Tamakloe T. Exploring Science & Mathematics Models through AI Technology to Advance High School Education Right. 2025. URL: https://www.researchgate.net/publication/390843042_Exploring_Science_Mathematics_Models_through_AI_Technology_to_Advance_High_School_Education_Right
9. Andrzejewski M, Dubicka N, Podolak J, Kowal M, Siłka J. Automated Test Generation Using Large Language Models. *Data*. 2025;10(10):156. doi:10.3390/data10100156
10. Круглова Н., Диховичний О. Застосування ІКТ у створенні програм-шаблонів для генерації тестових завдань з вищої математики. *Інформаційні технології і засоби навчання*. 2025. Том 108, № 4. С. 175–192. ISSN 2076-8184. DOI: <https://doi.org/10.33407/itlt.v108i4.6130>. URL: <https://journal.iitta.gov.ua/index.php/itlt/article/view/6130/2404>