

## ЧИ МОЖУТЬ ЧАТ-БОТИ ШТУЧНОГО ІНТЕЛЕКТУ СКЛАСТИ ІСПИТ З ФІЗИКИ?

Подласов Сергій Олександрович

старший викладач кафедри загальної фізики та моделювання фізичних процесів,  
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»  
[s.podlasov@kpi.ua](mailto:s.podlasov@kpi.ua)

**Постановка проблеми.** В останні два роки стрімко зросла кількість публікацій, присвячених чат-ботам штучного інтелекту (ChatbotsAI) в освіті. Автори публікацій або палко підтримують можливості та перспективи їх застосування, або ж критикують і називають «вбивцями освіти». Значно менше робіт присвячено дослідженню реальних можливостей ChatbotsAI, зокрема, їхній обізнаності у навчальних дисциплінах, які вивчають студенти технічного університету. Уявлення про це дозволить зробити висновок про їх спроможність надавати допомогу студентам при вивченні відповідних дисциплін.

Найбільш об'єктивною формою контролю обізнаності студентів вважається тестування. Так само тестування можна застосувати для перевірки обізнаності ChatbotsAI. Нами було поставлено завдання перевірити здатність вільно доступних ChatbotsAI (ChatGPT3.5, Copilot та Gemini) скласти іспит з фізики у тестовій формі і порівняти їхні результати з результатами тестового контролю студентів.

**Виклад основного матеріалу.** При навчанні фізики метою тестування є встановлення рівня оволодіння кожним конкретним студентом знаннями та уміннями відповідно до навчальних планів безвідносно до рівня підготовки інших студентів. Це означає, що тест є орієнтованим на критерій. Підготовка такого тесту передбачає складання специфікації тем, які підлягають контролю [1], відбору валідних завдань з урахуванням їхніх статистичних параметрів.

Для проведення поточного, тематичного та підсумкового контролю нами була підготовлена детальна специфікація і складені відповідні завдання, які розміщені в LMS Moodle. За результатами статистичного аналізу були відібрані валідні завдання трьох рівнів складності 1, 2, 3, які використовуються при проведенні підсумкового контролю (процедура аналізу та поділу за рівнями складності, а також процедура проведення іспиту детально описана в [2]).

Для з'ясування можливостей ChatbotsAI, їм були запропоновані ті ж самі завдання, які використовувалися при підсумковому контролі, набутих студентами, у літню екзаменаційну сесію 2022/23 навч. року та зимову сесію 2023/24 навч. року.

В зимову екзаменаційну сесію проводився контроль знань студентів за розділами «Механіка» та «Молекулярна фізика і термодинаміка». Всі студенти спочатку виконували завдання рівня 1 і могли перейти до виконання завдань рівня 2 якщо відсоток правильних відповідей був не менше 60 %. Так ж умова

була необхідна для переходу з рівня 2 на рівень 3. Із 103 студентів 78 змогли перейти з рівня 1 на рівень 2 і 44 студента виконували завдання рівня 3. Результати виконання завдань ChatGPT3.5, Copilot та Gemini показані в табл. 1, з якої очевидно, що на рівень 2 змогли б перейти тільки Copilot та Gemini (їхній підсумковій оцінці за 12-ти бальною шкалою можна співставити 6 – 7), оцінка ж GPT - незадовільна.

Таблиця 1.

Бали за тест рівня 1 зимової екзаменаційної сесії, набрані ChatbotsAI

Chatbot	GPT3.5	COPILOT	GEMINI
Відсоток правильно виконаних завдань	45	75	60

Незважаючи на незадовільні результати Chatbot3.5 усім чат-ботам були запропоновані завдання рівня 2. Сумарні бали за тестування (табл. 2) не дозволяють їм перейти до рівня 3.

Таблиця 2.

Бали за тест рівня 2 зимової екзаменаційної сесії, набрані ChatbotsAI

Chatbot	GPT3.5	COPILOT	GEMINI
Відсоток правильно виконаних завдань	32	50	45

В літню екзаменаційну сесію на іспиті перевірялися знання студентів з розділів «Електрика і магнетизм», «Коливання і хвилі» та «Елементи хвильової та квантової оптики». Завдання рівня 1 виконали 81 % студентів, що дозволило їм перейти до виконання завдань рівня 2. Ті ж самі завдання всі ChatbotsAI виконали незадовільно (див. табл. 3), що не дозволяє їм навіть претендувати на позитивну оцінку.

Таблиця 3.

Бали ChatbotsAI за тест рівня 1 весняної екзаменаційної сесії

Chatbot	GPT3.5	COPILOT	GEMINI
Відсоток правильно виконаних завдань	15	35	5

Однією з причин низьких показників CharbotsAI, може бути наявність в тестових завданнях рисунків з графіками, схемами, зображеннями векторів. Copilot та Gemini мають можливість аналізувати графічні дані, ChatGPT3.5 такої можливості не має. Як показали результати Copilot достатньо правильно інтерпретує близько 80 % графічного матеріалу і робить правильні висновки, у той час коли Gemini це робить приблизно у 40 % випадків.

Ще однією з причин надання невідних відповідей є тривіальні помилки в обчисленнях. Достатньо часто CharbotsAI наводять правильно виведені

остаточні формули, але обчислюють невірно. Оскільки більшість тестових завдань являє собою завдання множинного вибору або завдання з числовою відповіддю, то навіть при наявності правильної розрахункової формули, але неправильної числової відповіді завдання не вважається виконаним.

**Висновки.** Найвищий відсоток виконання тестових завдань показав чат-бот Copilot. Це може бути зумовлено тим, що в алгоритм його роботи закладений швидкий пошук правильних відповідей в мережі Інтернет, а також здатністю більш правильного розпізнавання та інтерпретації графічної інформації.

1) Випробування можливостей CharbotsAI скласти екзамен з фізики за програмою підготовки бакалаврів технічного напрямку, який проводиться у тестовій формі, показало, що їхні можливості достатні для одержання задовільної оцінки при здачі іспитів з розділів «Механіка» і «Молекулярна фізики та термодинаміка» і абсолютно недостатні для здачі іспиту за темами «Електрика та магнетизм», «Коливання та хвилі», «Елементи хвильової та квантової оптики». З цього випливає, що поки-що зарано говорити про широке застосування CharbotsAI для всебічної допомоги студентам при вивченні фізики.

Звичайно, зроблені висновки справедливі для сьогоденного рівня розвитку ШІ і можна сподіватися, що у подальшому всі ті сподівання, які на нього покладають, а також його власні претензії [3], будуть реалізовані.

### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. L. Crocker, J. Algina, Introduction to Classical and Modern Test Theory, Belmont, CA: Wadsworth, 527 p., 2006.
2. Подласов С. О., Матвійчук О. В. ПІДГОТОВКА ТА ПРОВЕДЕННЯ ЕКЗАМЕНУ З ФІЗИКИ В РЕЖИМІ ОНЛАЙН. Інформаційні технології і засоби навчання, 2022, Том 92, №6, с. 124-136.  
<https://journal.iitta.gov.ua/index.php/itlt/article/view/5068?articlesBySimilarityPage=2>  
DOI: 10.33407/itlt.v92i6.5068
3. S. O. Podlasov and O. V. Matviichuk, "APPLICATION OF CHATGPT IN THE TEACHING OF PHYSICS TO BACHELOR'S STUDENTS AT A TECHNICAL UNIVERSITY", ITLT, vol. 97, no. 5, pp. 149–166, Oct. 2023, doi: 10.33407/itlt.v97i5.5374.