

УДК 004.85:378.091.279.7]:004.655.3

DOI: <https://doi.org/10.30837/IYF.CVSAMM.2024.090>

**РОЗРОБКА МЕТОДУ АВТОМАТИЧНОГО ОЦІНЮВАННЯ
ВІДПОВІДЕЙ СТУДЕНТІВ З ВИКОРИСТАННЯМ
GPT-МОДЕЛЕЙ ВІД OPENAI ДЛЯ ВЕБ-ЗАСТОСУНКУ
З ТЕСТУВАННЯ ЗНАНЬ МОВИ SQL**

Науменко В.В.

Науковий керівник – доц. Яковлева О.В.

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

тел.: (057) 702-14-19, e-mail: vadym.naumenko@nure.ua

This paper is devoted to the development of a method for automatically evaluating students' answers using OpenAI GPT models for a web application designed to test SQL knowledge to replace the previous heuristic approach. The method uses the capabilities of GPT models to analyze a teacher's SQL script and a student's script, which allows to automatically evaluate students' answers when taking SQL tests. Thanks to this approach, the web application provides more efficient and automated testing, which will help bring the automatic system of evaluating students' answers closer to the level of evaluation by a real teacher.

На сьогодні суттєво збільшилася популярність платформ для онлайн навчання. Наприклад, в часи COVID попит на сервіси для проведення онлайн конференцій Zoom, Google Meet, MS Teams підвищився у декілька разів [1]. Але часто базової функціональності таких сервісів не вистачає для вирішення конкретних задач. Дана робота присвячена створенню веб-застосунку з тестування знань мови SQL, а саме розробці методу автоматичного оцінювання відповідей студентів з використанням GPT-моделей OpenAI. Існує багато онлайн-платформ для навчання програмуванню, проте доступ до найбільших сервісів, таких як Prometheus, зазвичай є платним. Серед безкоштовних ресурсів для перевірки знань з програмування варто відзначити CodeWars та LeetCode. Ці сервіси пропонують завдання різного рівня складності. Після успішного вирішення задачі користувачам доступне перегляд рішень, запропонованих іншими користувачами. Ця можливість сприяє обміну досвідом та пошуку оптимальних рішень. Проте ці платформи частіше використовуються людьми, в котрих вже є досвід в програмуванні і нажаль ці платформи не надають можливості структурованого навчання, а найбільшою проблемою є відсутність ідентифікації студентів. Саме тому було прийнято рішення розробити застосунок для навчання та тестування знань мови SQL, щоб використовувати його в університеті під час проведення дисциплін, пов'язаних з реляційними базами даних.

Головна відмінність цього застосунку від інших систем тестування полягає в тому, що застосунок даватиме змогу викладачеві організувати перевірку знань за необхідними темами, за заданим розкладом, а головне, сис-

тема буде дозволяти не обирати студентам правильні відповіді з переліку, а студент буде безпосередньо писати SQL скрипти, а система буде оцінювати правильність цих скриптів в автоматичному режимі. Система має підтримувати 3 типи користувачів: вчитель, адміністратор, студент.

Серверна частина нашого застосунку написана на .NET з використанням мікросервісної архітектури. Мікросервіси спілкуються один з одним за допомогою протоколу RPC, клієнт же в свою чергу написаний на фреймворку Angular. Дизайн клієнтської частини розроблений так, що усі користувачі мають однакову головну панель із кнопками авторизації та реєстрації та панель навігації (відрізняється тільки зміст даної панелі). Виходячи із вимог до застосунку певні типи користувачів мають деякі обмеження функціоналу. Усі ці обмеження впливають на те, що буде бачити перед собою користувач. Наприклад, користувач із роллю Адміністратор чи Вчитель має 7 пунктів меню на панелі навігації: Tests, Themes, Questions, Subjects, Users, Groups та Schedule Tests. Всі ці меню необхідні для створення, редагування та видалення, що дозволяє користувачу налаштувати питання, теми та тести так як йому потрібно. На основі сформованих запитань створюються тести, які студенти зможуть пройти та отримати оцінку. Для автоматичного оцінювання був запропонований евристичний алгоритм, що порівнює набори даних після виконання SQL запитів на основі еталонного скрипту та скрипту студента, а саме перевіряються співпадіння кількості стовпців та рядків, зміст комірок, використання необхідних операторів. Але такий підхід не перевіряє, чи реально скрипт студента виконує необхідне завдання. Може виникнути така ситуація, коли результати від еталонного скрипту та скрипту студента збігаються, але це випадковість і за наявності інших даних співпадіння не буде, тобто скрипт студента завдання не виконує.

Тому було прийнято рішення використовувати GPT модель від OpenAI, яка показала свою користь для вирішення багатьох задач [2], для покращення питання оцінювання відповідей студента. Для спілкування з моделлю GPT існують такі ролі: system – промпт, який допомагає моделі пояснити як треба виконати завдання; user – запит користувача; assistant – відповідь моделі. Для звертання до моделей OpenAI через API необхідно отримати ключ доступу на сайті OpenAI [3]. Нашим основним завданням є формування таких промптів, які допоможуть GPT моделі надати оцінку в балах та коментар-аргументацію, які будуть максимально схожі на оцінювання реальним викладачем. Приклад запиту та відповіді моделі GPT 3.5 turbo зображено на рисунках 1, 2.

На основі проведених досліджень можна зробити висновок, що GPT модель від OpenAI є потужною технологією для аналізу та оцінювання відповідей студентів. Вона надає докладний аналіз написаних студентом SQL скриптів та видає оцінку в балах. Використання GPT моделей допоможе максимально наблизити автоматичну систему оцінювання відповідей студентів до рівня оцінювання реальним викладачем. Також застосування

моделей GPT для пояснення помилок може бути використано не тільки для контролю знань студентів, а також під час вивчення мови SQL і підготовки до тестування.

Ти виконуєш автоматичне оцінювання знань студентів мови SQL ,
тобі дають:
1) структуру реляційної бази даних:
Product(IdProduct, NamtProduct,Price, Stock), де IdProduct - первинний ключ
Client(IdClient, NameClient, Address) , IdClient - primary key
Sale(IdSale, IdProduct, IdClient, Quantity, Date), ID_Sale - первинний ключ, IdProduct, IdClient -
foreing key.
2) завдання (що має скрипт отримати):
Вивести код та ім'я продуктів у яких ціна варіюються у проміжку між 300 та 500. У запиті
використати оператор Between
3) оператори, які треба використовувати або не використовувати студенту під час написання
скрипта:
треба використовувати BETWEEN;
4) еталонний (правильний)-скрипт:
SELECT IdProduct, NamtProduct
FROM Product
WHERE Price BETWEEN 300 AND 500;
5) скрипт-відповідь студента:
SELECT IdProduct, NamtProduct
FROM Product
WHERE Price >= 300 and Price <=500;
Тобі треба прокоментувати відповідь студента, наскільки вона правильна, чи відповідає вона
завданню і поставити оцінку в діапазоні від 0 до 1 за таким принципом:
- оцінка 1, якщо в результаті виконання скрипта студента:
а) буде виконуватися поставлене завдання (пункт 1);
б) у результаті еталонного скрипта і скрипта студента вийде однаковий набір кортежів
(збігається кількість рядків і стовпців, а значення клітинок збігатимуться);
в) буде використано необхідний оператор (пункт 3);
- оцінка 0.7:
а) буде виконуватися поставлене завдання (пункт 1);
б) у результаті еталонного скрипта і скрипта студента вийде однаковий набір кортежів
(збігається кількість рядків і стовпців, а значення клітинок збігатимуться);
в) буде НЕ використано необхідний оператор (пункт 3);
- оцінка 0: якщо в результаті виконання еталонного скрипта і скрипта студента НЕ вийде
однаковий набір кортежів
Постав, будь ласка, оцінку в бальній системі від 0 до 1 і дай коментар до відповіді студента
щодо правильності та відповідності завданню

Рисунок 1 – Приклад промπτу

Відповідь студента в цілому вірна, оскільки вона виконує поставлене завдання і видає очікувані результати. Однак, студент використав два оператори порівняння (>= і <=), коли було вимога використовувати оператор BETWEEN.

Оператор BETWEEN використовується для визначення діапазону значень у вибірці. Він дозволяє спростити запит, який відповідає на питання про значення, що потрапляють у певний діапазон. У цьому випадку, оператор BETWEEN є більш зрозумілим та коротким рішенням.

Оцінка: 0.7

Рисунок 2 – Відповідь моделі на запит з рис. 1

Список використаних джерел:

1. Yakovleva, O., Kovač, M., Ardasov, V. & Yeremenko, I. (2023). Study on adding functionality to the Zoom online conference system for monitoring the participant activities. *Public Administration and Regional Development*, 19(1), pp. 158–184.

2. Yakovleva O., Nebeský L., Kirichenko A. (2023) Using the GPT models for responses based on custom content to develop neural consultant for university applicants. *Abstracts of V International Scientific and Practical Conference*. Madrid, Spain. pp. 172–178.