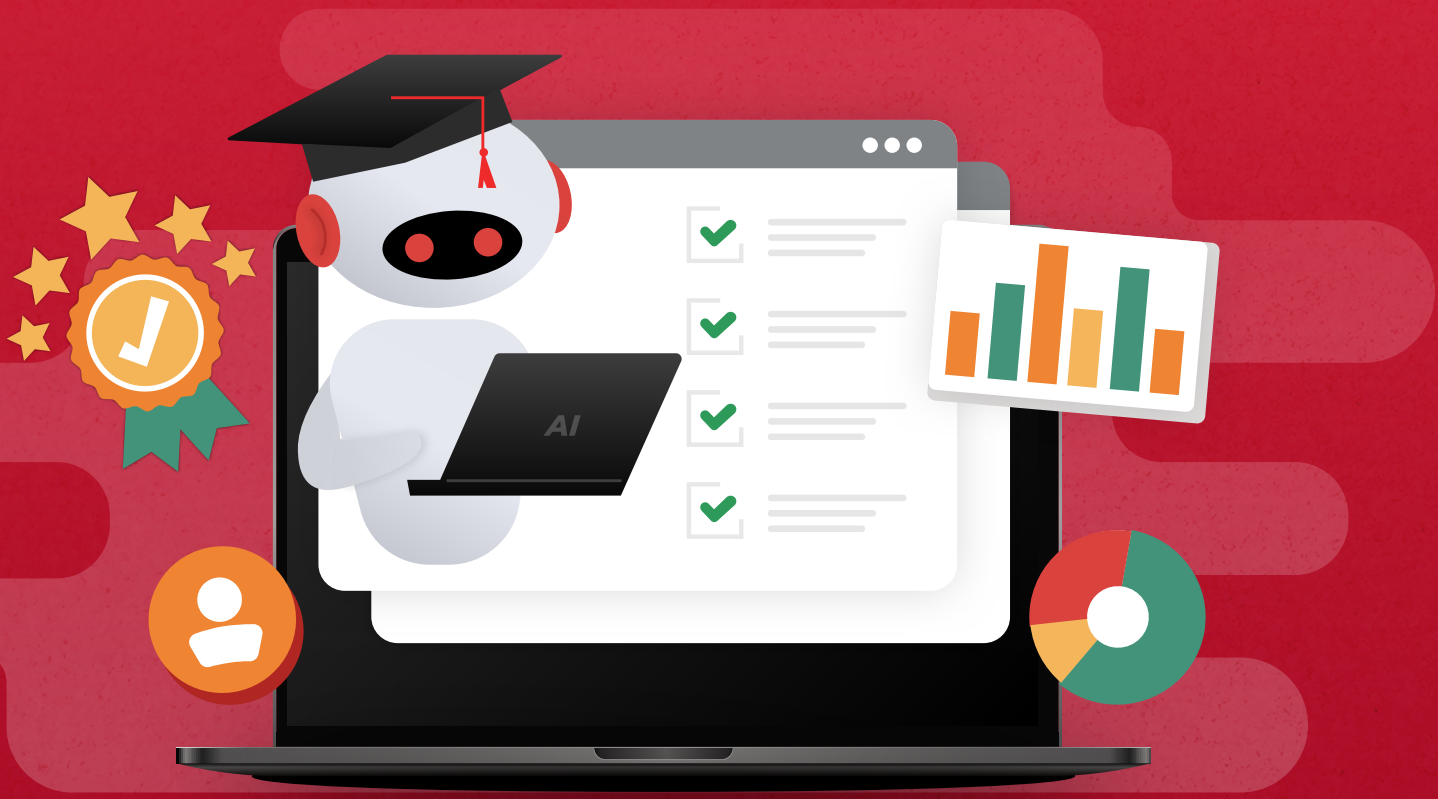
QBurst

AI-Automated Student Grading System

By leveraging Large Language Models (LLMs) and tailored prompt engineering techniques, we developed an efficient and cost-effective solution for automating the evaluation of student test responses.



Overview

Using advanced language models, we automated the evaluation of student test responses to achieve accurate and consistent grading. The solution is scalable, easily adapts to different educational environments, and requires minimal manual intervention.

We utilized OpenAI's GPT-3.5 and GPT-4 models, along with custom prompt engineering techniques, to develop a robust evaluation system. The system analyzes and grades student responses, offering comprehensive feedback while maintaining high accuracy. We also conducted a comparative study to validate the system's performance against grades assigned by experienced educators and found a high level of alignment in grading outcomes.

Client Profile

Headquartered in the United States, our client is a leading provider of educational technology platforms designed to help students enhance their writing and grammar skills. The platform focuses on providing personalized, adaptive learning experiences for K–12 students through engaging and relevant content.

Business Requirements

The client wanted to implement an automated solution for evaluating student test responses. Their goal was to develop a comprehensive evaluation system that would provide accurate and consistent grading across diverse educational environments.

To achieve this, they sought a unified solution leveraging advanced AI techniques that would automate grading, alleviate teacher workloads, and offer prompt feedback to students.

QBurst Solution

We developed an automated grading solution using OpenAI's advanced language models, GPT-3.5 and GPT-4, to accurately and efficiently evaluate student test responses. Despite initial considerations to use open-source LLMs to reduce costs, we chose OpenAI's models, largely due to its superior accuracy and reliability.

To meet business requirements, we developed a solution that serves as a single package for evaluating student answers, assigning grades, and providing individual feedback. The system employs custom prompt engineering techniques to evaluate student responses in parts. This modular approach allowed us to methodically evaluate various aspects of the answers, ensuring a thorough and detailed evaluation. The Gradio app shows the final grade with grade breakdown.

Technologies Used

python™ gradio OpenAI pytest

Business Benefits

- **Improved Productivity**

  The AI-driven evaluation platform significantly increased productivity by allowing educators to focus more on teaching and less on the time-consuming grading task.

- **Enhanced Efficiency**

  Automating the evaluation process reduced the grading time dramatically, enabling teachers to allocate more time for student engagements and curriculum development.

- **Consistent and Fair Grading**

  The use of advanced language models ensures standardized and unbiased grading by providing uniform assessments for all students.

- **Detailed Feedback**

  The solution provides comprehensive feedback on each student, helping them understand their mistakes and areas for improvement.

- **Scalability**

  The system, designed to be easily scalable, can accommodate a growing number of student responses, student groups, and programs.

- **Cost-efficiency**

  The investment in OpenAI's models proved economical due to the high level of accuracy and reduced need for manual intervention.