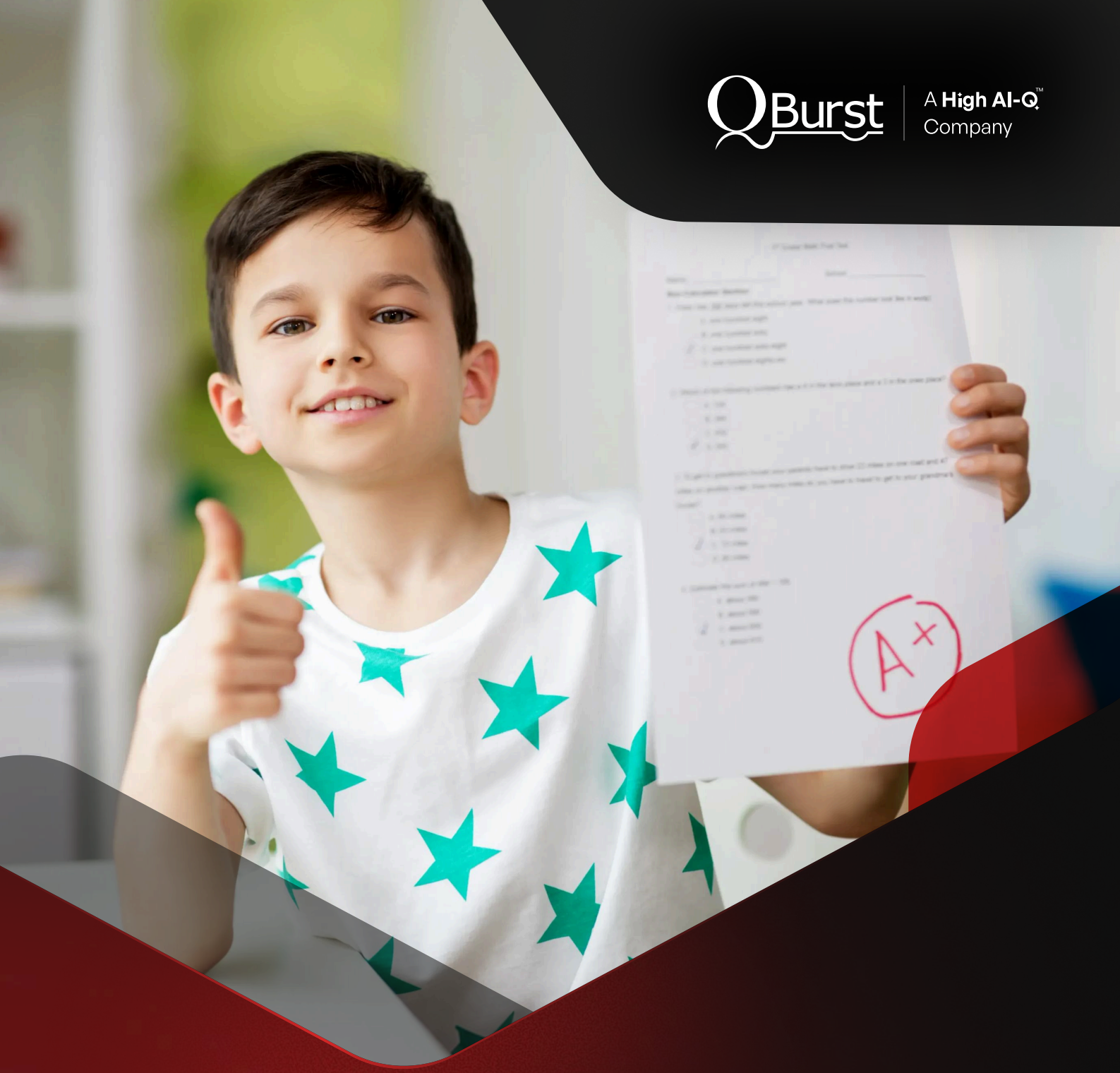
QBurst | A High AI-Q™ Company



AI-Automated Student Grading System for Streamlined Evaluations

A unified, accurate, and scalable automated grading system using GPT-4 and custom prompt engineering to significantly reduce teacher workload for a leading US EdTech platform.

Overview

A leading US-based EdTech provider sought to leverage Large Language Models (LLMs) to automate the time-consuming and labor-intensive process of grading student test responses.

- Delivered a robust, single-package automated grading solution utilizing OpenAI's GPT-3.5 and GPT-4, ensuring high accuracy and reliability.
- Developed custom prompt engineering techniques to break down and methodically evaluate various aspects of student answers, providing detailed, modular feedback and a grade breakdown via a Gradio app.
- Significantly enhanced productivity and efficiency for educators by automating assessment, leading to faster feedback for students and consistent, unbiased grading.

Client

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.

Challenges: Inconsistent Grading and Teacher Workload

The client's need for modernization centered on addressing key operational bottlenecks.

- **High Teacher Workload:** Manual grading of student responses was a time-consuming task, diverting educator focus from teaching and curriculum development.

- **Inconsistent Grading:** Maintaining consistent and unbiased grading standards across a large, diverse student base and varied educational environments was a challenge.
- **Slow Feedback Cycle:** The time required for manual grading delayed feedback to students, impeding their adaptive learning process.
- **Need for Scalability:** The evaluation system needed to easily accommodate a growing volume of student responses, student groups, and programs.

QBurst Solution: LLM-Driven Evaluation Platform

We developed an AI-automated grading solution that serves as a single, unified package for evaluating student answers, assigning grades, and providing individualized feedback. The system is built on OpenAI's GPT-3.5 and GPT-4 models, chosen for their superior accuracy and reliability over initially considered open-source alternatives.

The core innovation lies in the custom prompt engineering techniques employed. The solution evaluates student responses in modular parts, breaking down the assessment criteria to ensure a thorough and detailed evaluation of various aspects of the answers.

- The system analyzes the student response against the correct answer and grading rubrics defined in the prompts.
- A user-friendly Gradio application was built to display the final grade along with a detailed grade breakdown, enhancing transparency and user experience.
- PyTest was used to ensure the reliability and correctness of the underlying Python code.

Technical Highlights

- Leveraged OpenAI's GPT-4 for its advanced natural language understanding and reasoning capabilities.
- Implemented modular prompt engineering for multi-faceted evaluation of student responses.

- Conducted a comparative validation study against experienced educators' grades, confirming a high level of alignment and accuracy.
- Designed the system for easy scalability to handle a rapidly growing number of users and programs.

Impact: Automating Evaluations with Speed and Accuracy

The AI-Automated Student Grading System delivered measurable benefits to the client's educational ecosystem.

- **90%+ Productivity Gain:** Educators can now focus significantly more time on teaching and curriculum development rather than grading.
- **Consistent and Fair Grading:** The use of advanced LLMs ensures standardized and unbiased assessment across all students.
- **Accelerated Feedback Cycle:** Automated grading dramatically reduces the time between test submission and feedback receipt.
- **Detailed Feedback:** Students receive comprehensive, individualized feedback, helping them to quickly identify and address areas for improvement.
- **Cost-Efficiency:** The high accuracy and reduced need for manual intervention made the investment in GPT models economical and valuable.