A Comprehensive Analysis of the AI Assessment Tool Gradescope (by Turnitin)

Description

Introduction

The rise of artificial intelligence (AI) has revolutionized many fields, including education. Gradescope was developed in 2014 by UC Berkeley and joined Turnitin, LLC, a platform for detecting plagiarism, in 2018 (Turnitin. n.d.). It is an AI-powered assessment tool designed to streamline grading processes, improve feedback quality, and enhance studentsâ?? learning experiences through formative assessment features (Standford Centers for Teaching and Learning). This paper provides a detailed analysis of Gradescope, including its features, usage instructions, benefits, challenges, and perspectives from teachers from survey results. Additionally, it evaluates the toolâ??s effectiveness, biases, reporting mechanisms, and appropriateness, offering suggestions for improvement.

default watermark

Purpose of the Analysis

According to the Gradescope website (www.gradescope.com), more than 700 million questions have been graded by 2,600 Universities using Gradescope and 140k instructors, and 3.2 million students have used the platform. Gradescope may have been user-approved over time, backed by its widespread usage and adoption by major universities nationwide. However, the goal of this paper is not to necessarily recommend the tool but to critically review it.

Gradescope: Type of Assessment

(Retrieved from website www.gradescope.com)

1. Homework Assignments: Students can submit handwritten or typed assignments, which Gradescope digitizes and organizes for grading.

2. Exams: Instructors can create and distribute exams, which Gradescope can then grade based on pre-defined rubrics. Automatically grade versioned, multiple-choice exams, complete with advanced statistics.

3. Quizzes: Similar to exams, AI can help grade quizzes automatically or manually.

4. Programming Assignments: Gradescope supports automatic grading of code submissions, providing immediate feedback to students.

5. Grade All Subjects: Gradescope supports variable-length assignments (problem sets and projects) and fixed-template assignments (worksheets, quizzes, bubble sheets, and exams).

Key Features

1. AI-Assisted Grading: Automates grading by recognizing patterns and applying consistent grading criteria.

2. Rubric-Based Grading: Allows instructors to create detailed rubrics that ensure uniform grading standards.

3. Feedback Mechanisms: Provide detailed feedback to students, highlighting areas of improvement. Leading-edge AI technology groups similar student answers, eliminating redundancies and ensuring consistent feedback.

4. Statistical Analysis: This offers insights into class performance and identifies trends or common mistakes. Aggregate data informs department-level metrics. Detailed Analytics per-question and per-rubric statistics help understand how students are doing.

5. Integration with Learning Management Systems (LMS): Seamlessly integrates with popular LMS platforms like Canvas and Blackboard.

6. Support for Multiple File Types: Accepts various file formats, including PDFs, images of handwritten works, and programming code.

7. Integrity: Gradescope mitigates risks related to integrity, such as copying or bias in grading.

default watermark

8. Accessing Help: Gradescopeâ??s built-in Knowledge Bot provides immediate access to help articles without leaving the platform.

9. Student Mobile App: Students can conveniently upload assignments using the Gradescope mobile app.

10. Gradescope Roadmap: Customers can submit idea cards for desired features and vote on features that are being considered. They can also create a Known Issue card to report issues while using the platform and let others view it.

11. Multilanguage: Fully operates in English, Japanese, Korean, Spanish, and Turkish.

Instructions for Use

1. Setup: Instructors sign up for an account and create a course. They can then add assignments, exams, or quizzes.

2. Submission: Students submit their work through Gradescopeâ??s portal. The tool supports both individual and group submissions.

3. Grading: Instructors create rubrics and use Gradescopeâ??s AI to assist in grading. The tool can automatically grade multiple-choice and short-answer questions and assist with longer responses.

4. Feedback: Once grading is complete, instructors can release grades and feedback to students through the platform.

5. Review and Analysis: Instructors can review statistical data on student performance to identify areas that may need additional attention.

Results from Research Findings:

Unfortunately, there wasnâ??t much data available from external research studies. The findings I referenced are from a 2017 Gradescope and UC Berkeley study, â??Gradescope: a Fast, Flexible, and Fair System for Scalable Assessment of Handwritten Work.â?• The study was conducted to measure its effectiveness in three key dimensions: speed, consistency, and flexibility. The results are based on user-reported data and survey feedback regarding time-saved grading, user-friendliness, and student experience.

Speed and Grading Efficiency

Decreasing Time Per Submission: The time spent grading each submission decreases as more submissions are graded because of the rubric setting feature.

Time Savings: A survey revealed that 67% of users reported saving 30% or more time compared to traditional paper-based grading. This efficiency is critical in large courses where timely feedback is essential.

Consistency and Fairness

default watermark

Consistency: Using a dynamic rubric ensures that all students are graded according to the same standards, improving grading fairness. Most users reported that the system helps them grade more fairly, with 46.4% strongly agreeing and 33.3% agreeing.

Inter-Grader Reliability: Rubric-based grading increased inter-grader reliability, which benefited courses with multiple graders. Additionally, anonymous grading redacts student names and IDs during assessment viewing, reducing unintended bias.

Flexibility and Feedback

Dynamic Rubrics: The dynamic rubrics allow graders to add new types of mistakes as they encounter them and retroactively update all previous grades accordingly. This flexibility helps maintain consistency and allows for more detailed feedback.

Detailed Feedback: Students receive detailed feedback on their mistakes, which helps them learn and improve. 45.6% of instructors agreed that the rubric displayed helped students learn more from their mistakes.

Handling Regrade Requests

Simplified Regrade Process: The system tracks which grader graded each studentâ??s answer and notifies the grader when a regrade is requested. This process is more efficient than traditional methods, reducing turnaround time for regrade requests. 41.2% of users strongly agreed that the system simplified the regrade request process.

Analysis of Student Performance

Rubric-Level Statistics: The system tracks detailed statistics on student performance, including which mistakes were most common. This data is invaluable for instructors to identify common misconceptions and adjust their teaching accordingly.

Insight into Misconceptions: Analyzing rubric-level data allows instructors to gain insights into specific areas where students struggle, enabling targeted interventions.

Usage Statistics

Scalability: Gradescope analyzed databases from over 200 institutions and 10 million pages of student work, demonstrating its scalability and wide adoption in diverse educational settings.

Variety of Assignments: The system supports various assignments, from short quizzes to comprehensive exams, highlighting flexibility.

Survey Results

Fair Grading: Does the system help you grade more fairly?

Strongly Agree: 46.4%

Agree: 33.3%

default watermark

Time Efficiency

Does the system save you time in grading?

Strongly Agree: 60.9%

Agree: 26.1%

How much time do you save grading with our system versus grading on paper?

10%: 91%

20%: 88%

30%: 67%

Enjoyable Grading: Does the system make grading more enjoyable?

Strongly Agree: 26.9%

Agree: 32.8%

Transparency: Does the system offer transparency to my students about the grading scheme?

Strongly Agree: 38.2%

Agree: 44.1%

Learning from Mistakes: Does the displayed rubric help your students learn more from their mistakes?

Strongly Agree: 17.6%

Agree: 45.6%

Results from another study by Turnitin, covering Gradescope data collected between August 2013 and September 2018, a dataset of 242,775 graded student responses across 1,358 questions, 338 courses, and 43 institutions, also showed similar positive findings regarding assessment accuracy and time efficiency (Yen et al., 2020).

Challenges

Technology Literacy & Training

Teachers and students may face a learning curve when using the Gradescope platform for the first time. Some educators find the initial setup and training to be time-consuming. Many reported initial difficulties in understanding and navigating the platform, which impacted their ability to utilize it fully in the early stages (Hall, 2023). This finding highlights the need for comprehensive training and support for new users to ensure smooth adoption and effective tool use (Tsai, 2024).

Privacy

default watermark

Under the â??Terms of Use,â?• there was some concerning wording regarding student data. Turnitin ensures that the data is de-identified to avoid identifying specific schools or individuals (Tsai, 2024). However, it is stated that students grant Turnitin a non-exclusive, royalty-free, worldwide license to use their data. More measures must be taken to protect individual privacy and hold accountability to ensure compliance with legal requirements. Dealing with studentsâ?? performance data and profiling based on grades is risky. Gradescope should actively communicate with stakeholders regarding risks and possible issues in privacy.

Use of Student Data

By submitting Student Data or other information to the Gradescope service, you expressly grant, and you represent and warrant that you have all rights necessary to grant, to Turnitin a non-exclusive, royalty-free, worldwide license during the term of these Terms to use, transmit, distribute, modify, reproduce, display, create derivative works of, and store the Student Data solely for the purposes of (i) providing the Services as contemplated in these Terms, and (ii) enforcing its rights under these Terms.

Anonymized Data

You agree that Turnitin may collect, analyze, and use data derived from Student Data, including de-identified, aggregated or anonymized Student Data, as well as data about your, and other usersâ?? access and use of the Service, for purposes of operating, analyzing, improving, or marketing the Service and for the purpose of providing analytic services to the School or to other Schools. If Turnitin shares or publicly discloses information (e.g., in marketing materials, in application development, or with third parties) that is derived from Student Data, such data will be de-identified to reasonably avoid identification of a specific school or individual. You further agree that Turnitin will have the right, both during and after the Term of these Terms, to use, store, transmit, distribute, modify, copy, display, sublicense, and create derivative works of the de-identified, aggregated or anonymized data solely for the purposes of improving Turnitin services.

(Retrieved from website www.gradescope.com)

Digital Equity

Reliance on AI and technology may exclude students with limited access to digital resources. A study on the digital divide in education, surveying students from various socioeconomic backgrounds, found that students from lower-income families often lacked reliable internet access or up-to-date devices, limiting their ability to benefit from tools like Gradescope (Gonzalez, 2023). The finding underscores the importance of ensuring equitable access to digital resources to avoid exacerbating educational inequalities.

Bias and Fairness

Even though rubric-based grading ensures consistent criteria are applied to all students (Singh et al., 2017). Concerns have been raised about potential biases in AI algorithms. They found that while Gradescope generally performs well, there are instances where AI could inadvertently favor certain types of responses over others (Reck, 2019). Gradescope claims the platform is designed to recognize patterns without bias and to prevent unintended biases (Retrieved from Gradescope website). Continuous monitoring and updates to the algorithms are necessary to mitigate these biases and ensure fair assessments for all students.

default watermark

Lack of Independent Research

Independent research ensures that evaluations are unbiased and not influenced by corporate interests. Corporate data may be selectively shared or manipulated to present a favorable view of the tool, whereas independent research aims for objectivity (Gonzalez, 2023). More accurate data and statistical insights into Gradescopeâ??s impact are necessary for improvement.

Suggestions for Improvement

- Enhanced Training: Provide more comprehensive training resources for teachers and students. Provide more streamlined onboarding for educators with limited technology literacy. Provide ongoing PD to assist educators in technology literacy and using the features fully.
- Equity and Accessibility: Ensure the platform is accessible to all students, regardless of their technological resources. As a corporation profiting from the public sector, it should invest more money to make the platform equitable and accessible to all students from diverse socioeconomic backgrounds.
- Continuous Monitoring and Transparency in Privacy: Regularly update AI algorithms to prevent and address emerging biases. Since Gradescope is widely used by secondary education institutes, monitoring or auditing from a third-party entity or stakeholder like the Department of Education is needed. More importantly, transparency about privacy and the use of student data is also needed.
- Independent Research: Independent researchers can scrutinize AI tools more rigorously. They can assess transparency (availability of source code, algorithms, and data) and hold developers accountable for any shortcomings. Gradescope needs to back up its claims about its effectiveness with more independent studies.

Conclusion

Overall reviews and feedback from educators indicate that Gradescope is a fast, flexible, and effective system for grading and assessment (Tsai, 2024). Even though the findings were from studies conducted by the company that owns Gradescope, the research data was credible because of its substantial sample size. A larger sample provides more reliable estimates and reduces the impact of

random variation. With a sizable dataset, the findings are more likely to represent the broader population, enhancing the studyâ??s validity (Arthurs et al., 2019). The findings showed significant time savings, increased consistency, detailed feedback, and efficient handling of regrade requests, highlighting its effectiveness (Singh et al., 2017). The scalability and wide adoption of Gradescope further reinforce its utility in current higher education (Standford Centers for Teaching and Learning).

According to a survey of 924 educators conducted by the EdWeek Research Center in 2023, nearly 80 percent of educators say their district has not crafted clear policies on using AI in the classroom (Klein, 2024). The Peninsula School District in Washington state was among the first school systems in the country to put out guidance on using artificial intelligence (AI) in the classroom (Klein, 2024). The University of Washington (UW) has developed a product called Colleague, which uses AI and chatbots to assist K-12 teachers in creating lesson plans (Stiffler, 2024). We need to assess AI tools critically and provide transparent data to educators so that they can make informed decisions. Also, district administration must use evidence-based data for decision-making on implementing AI tools and develop comprehensive guidelines for using AI tools. While there are challenges and areas for improvement, the benefits of Gradescope make it a valuable tool for assessment. The best feature is that it allows teachers to use their own rubrics and lets teachers have control over how the grading is done. Having a human-centered approach is significant in managing AI tools, as the Office of Superintendent of Public Instruction (OSPI) AI guideline recommends (OSPI: https://ospi.k12.wa.us/student-success/resources-subject-area/human-centered-artificial-intelligence-schools).

default watermark

References

Arthurs, N., Stenhaug, B., Karayev, S., & Piech, C. (2019). Grades Are Not Normal: Improving Exam Score Models Using the Logit-Normal Distribution. International Educational Data Mining Society.

Hall, E., Seyam, M., Dunlap, D. (2023). Identifying Usability Challenges in AI-Based Essay Grading Tools. Communications in Computer and Information Science, vol 1831. Springer, Cham. https://doi.org/10.1007/978-3-031-36336-8_104

Gonzalez, V. H., Mattingly, S., Wilhelm, J., & Hemingson, D. (2023). Using artificial intelligence to grade practical laboratory examinations: Sacrificing studentsâ?? learning experiences for saving time?. Anatomical Sciences Education.

Klein, A. (2024, February). Need an AI Policy for Your Schools? This District Used ChatGPT to Craft One. EdWeek. https://www.edweek.org/technology/need-an-ai-policy-for-your-schools-this-district-used-chatgpt-to-craft-one/2024/02

Reck, R. M. (2019). A systematic review of technologies for providing feedback and grades to students. ASEE Annual Conference and Exposition, Conference Proceedings. https://doi.org/10.18260/1-2â??32008

Singh, A., Karayev, S., Gutowski, K., & Abbeel, P. (2017). Gradescope: a fast, flexible, and fair system for scalable assessment of handwritten work. In Proceedings of the fourth ACM conference on learning@ scale (pp. 81â??88).

Standford Centers for Teaching and Learning: https://ctl.stanford.edu/use-learning-technology/gradescope

Stiffler, L. (2024, May 24). Chatbots for teachers: Univ. of Washington releases free AI tool for quicker, better lesson plans. GeekWire

Tsai, T. (2024, April 18). Gradescope Review: Is It Worth It? SCI Journal. https://www.scijournal.org/articles/gradescope-review

Turnitin. (n.d.). Gradescope | Modern Grading & Assessment Platform.

Turnitin (2019, Sept 17). Gradescope Increases Grading Consistency and Student Engagement in Computer Science Courses.

Yen, M., Karayev, S., & Wang, E. (2020). Analysis of grading times of short answer questions. In Proceedings of the Seventh ACM Conference on Learning@ Scale (pp. 365-368).



Category

1. Uncategorized

default watermark

Date Created
November 6, 2024
Author
admin