

# AI-Driven Exam Grading, Percentile Calculation, and Plagiarism Detection System for Educational Integrity

Tajjuddin Malnas<sup>1</sup>, Prasanna Pandhare<sup>2\*</sup>, Divyank Panwar<sup>3</sup>, Sphurti Deshmukh<sup>4</sup>

## Abstract

*The influence of artificial intelligence (AI) on academic evaluation has become increasingly significant, particularly in areas such as automated grading, large-scale content analysis, and plagiarism detection. These technologies provide educators with rapid processing, consistent scoring, and timely feedback that can support both teaching and learning. However, alongside these advantages, serious concerns have emerged regarding the reliability and fairness of AI-driven assessments. The growing presence of AI-generated content has further complicated the evaluation process, often leading to false positives, misclassification, and questionable judgments about academic integrity. These issues highlight the ethical challenges educators face when relying solely on automated systems. Therefore, a more balanced and responsible approach is essential—one that blends human expertise with the analytical power of AI. Such a hybrid framework can help reduce errors, improve transparency, and maintain accountability within academic environments. By combining human judgment with technological efficiency, institutions can promote a more reliable, ethical, and high-quality assessment process, ensuring that AI-supported educational tools genuinely enhance learning rather than compromise it.*

**Keywords:** AI-powered education, grading, academic integrity, plagiarism detection, hybrid framework, ethical education, AI content, human-AI collaboration, content analysis, human insights, misclassification

## INTRODUCTION

Currently, artificial intelligence (AI) is used in academic evaluation, content verification, and validation. It can also be used for plagiarism detection and AI content recognition. These technologies help save time and provide constant feedback. However, various issues arise owing to complex AI models, which can be inaccurate and can lead to misclassification. Research has shown that the combination of AI with human monitoring reduces danger while retaining educational integrity and fairness. Issues such as privacy, ethics, and potential biases have been raised in automated grading systems. This suggests the importance of human judgment in a high-stakes context. The use of AI in education is similar to that in other sectors, where a hybrid framework balances efficiency and reliability. Progress in AI has helped sustain ethical standards and maintain quality feedback, which has ensured the retention of academic confidence and transparency.

### \*Author for Correspondence

Prasanna Pandhare

E-mail: [prasannapandhare20160@gmail.com](mailto:prasannapandhare20160@gmail.com)

<sup>1-3</sup>Student, Department Information Technology, Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India

<sup>4</sup>Assistant Professor, Department Information Technology, Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India

Received Date: May 19, 2025

Accepted Date: September 16, 2025

Published Date: November 17, 2025

**Citation:** Tajjuddin Malnas, Prasanna Pandhare, Divyank Panwar, Sphurti Deshmukh. AI-Driven Exam Grading, Percentile Calculation, and Plagiarism Detection System for Educational Integrity. *Current Trends in Information Technology*. 2026; 16(1): 7–12p.

## LITERATURE SURVEY

The study highlights the merits and limitations of AI content identification technology [1]. This can be achieved by distinguishing between AI-generated and human-authored materials. Compared to

---

previous AI models, these tools are more accurate. However, these automated detections can lead to misclassifications, which can affect academic performance. The reason for misclassification is the presence of false positives in human-written content. Current advancements in AI suggest a combination of human and AI models to create a hybrid framework. Such improvements help keep up with the increasing sophistication of AI. The study helps understand the use of AI in evaluating student papers [2]. Along with this evaluation, it can suggest some of the benefits and drawbacks. Key advantages include timesaving, convenience, and continuous feedback. This can help to improve the learning process. Drawbacks can be related to privacy, legal consequences, and biases in the AI-generated content. The study also stated that relying on AI for complete grading can be hazardous. A balanced hybrid framework can help to incorporate academic integrity and quality feedback. The study introduced the concept of Gen-AI for grading open-ended responses [3]. This model can evaluate responses by combining large language models, sample solutions, and grading parameters. It is an open-source tool that allows flexibility, customizable grading, time-saving, and improved uniformity. Expert ratings, good reliability, and accuracy were verified measures for educational assessments. This program has limitations, such as complicated questions and ethical concerns regarding the use of AI for graded examinations.

The study shows how students use AI in writing academic essays and highlights its benefits and drawbacks [4]. Writing skills, grammatical checks, self-efficacy in critical thinking, and plagiarism detection can be enhanced using AI tools. This study also shows how AI can affect students' originality, creativity, and critical thinking. This study recommends the use of AI sensibly to improve writing without sacrificing students' critical thinking skills. The findings of this study suggest combining AI assistance with human input. This can increase productivity while preserving academic integrity. To successfully analyze academic writing tasks, the study introduced plagiarism detection tools that can be examined through various files and folders [5]. Cosine similarity can be used to analyze the degree of plagiarism. These tools can handle various issues on their own, such as student-to-student copying. It also offers a more efficient solution than classic scanners. The last study stated that future advancements will use Optical Character Recognition (OCR). OCR is an image-based plagiarism detection method. This can help improve academic integrity. The study highlighted a teaching strategy that can help prevent plagiarism [6]. Plagiarism decreased by encouraging comprehension of academic integrity and enhancing students' writing abilities. This plagiarism decreased by more than 37% and was implemented as part of English for academic purposes. This technology can effectively identify contract cheating. However, there are some limitations to distinguishing between deliberate and unintentional misconduct. This demonstrates how important it is for students and teachers to continue their education.

Elkhatat et al. investigated various ways in which students can avoid text-matching technologies for plagiarism detection [7]. Text conversion to pictures, adding invisible letters, and replacing characters with Unicode symbols are some of the strategies that can be used. Nine tools were tested to determine their usefulness. Among these, the majority of systems can detect conventional plagiarism but struggle with more advanced techniques. This study states that detection reliability can be improved by using OCR and editable file formats. This study provides remarkable insights [8]. For example, 15 web-based text-matching tools in eight different languages were tested against the efficiency of identifying plagiarism. Plagiarism can be found in single- and multi-source publications. Usability, source type recognition, and language coverage are some of the metrics used to identify the performance. This experiment showed that many systems have trouble with translation, synonym replacement, and paraphrasing. However, some are good at detecting straight copying. To achieve academic standards, studies have revealed differences in language support, report clarity, and simple use, which indicates the need for enhanced features. The study showed the limitations of conventional evaluation techniques [9]. This also shows how AI can improve assessment by handling problems such as scalability, engagement, and adaptability. AI is used in adaptive testing, essay scoring, peer evaluations, and continuous performance monitoring to reduce teacher workload and promote individualized instruction. Nonetheless, the writers also draw attention to difficulties, such as possible prejudices, a reduction in teacher autonomy, and moral dilemmas. According to the report, maintaining educational quality and accountability requires striking a balance between AI and human control.

With emphasis on text rewriting and grammar correction, Chemaya and Martin [10] explored academic perceptions and the detection of AI use in manuscript writing. The results indicate that most academics do not feel compelled to reveal few grammatical modifications, and there is disagreement over whether or not to acknowledge significant text rewrites. This study used a detection technique to find discrepancies because even minor grammatical changes resulted in high AI scores. According to the research, to harmonize community norms and detection methods, precise criteria for the use of AI in academic writing are required. The effectiveness of machine learning in lowering the workload of teachers who grade short answers was assessed [11]. Up to 74% less manual grading was required when machine learning and manual grading were combined. A classifier is developed to grade the remaining responses. Grouping-related responses made the training and final grading procedures run smoothly. Although this method offers a scalable solution for large courses, accuracy and fairness still require manual assessment of automated marks. With an emphasis on techniques such as oral exams, collaborative projects, and conceptual assignments that restrict AI's advantage, Xie et al. [12] explored the ways to lessen AI-assisted cheating. This study emphasizes the importance of creating AI tools for identifying AI-generated materials, incorporating ethics education, and promoting industry-academia partnerships. Together, these initiatives tackle ethical and technical issues, with the goal of maintaining educational integrity in the face of the growing influence of generative AI.

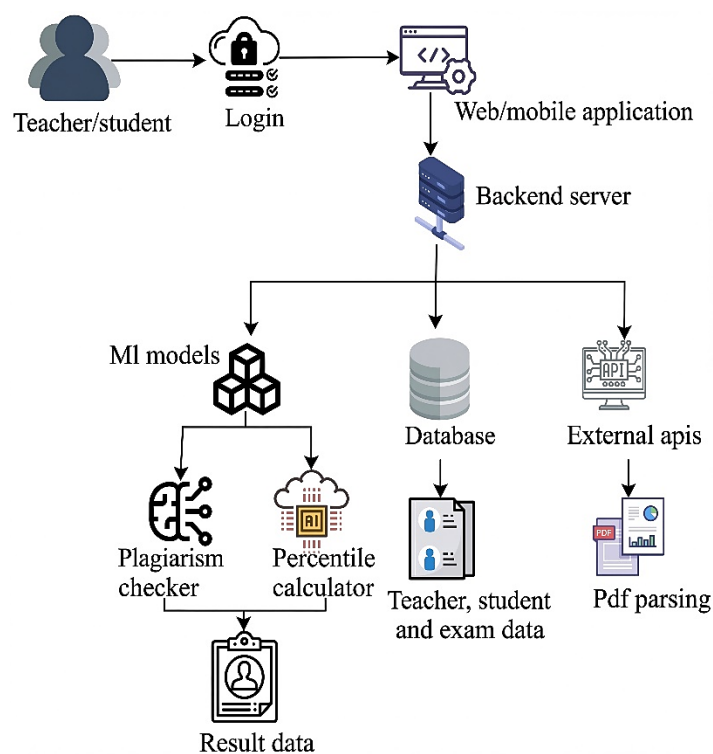
To improve grading efficacy and cost-effectiveness, Yimyam and Ketcham [13] presented a multiple-choice exam grading system that uses mobile phone image processing. The system achieved 96.87% accuracy on 1,600 items by identifying answers from the collected photos using edge detection and pixel intensity calculations. Although restricted to particular answer sheet formats, this mobile-based method is a useful and approachable substitute for conventional grading methods, making it perfect for environments with minimal resources. An e-assessment system that uses image processing to assess multiple-choice answers has also been presented [14]. Instead of employing expensive Optical Mark Recognition (OMR) machines, the system processes answer sheets digitally by utilizing OpenCV within the Django framework. This system facilitates graphical performance visualization, outcome analysis, and automatic assessment. Although it offers a scalable and economical grading option, nonobjective question categories have not yet been supported. To make online grading easier, Migut and Wiersma [15] suggested an effective way to digitize paper tests. A video-based pipeline is used in this process, which recognizes important exam page frames, extracts and fixes page perspectives, associates questions with pages, and recognizes student information using character recognition. By performing manual scanning, increasing grading efficiency, and integrating with digital grading platforms, this technology makes remote, secure, and readable grading possible. In the future, character recognition will be improved to achieve greater accuracy.

Cao et al. investigated the use of image processing to assess mathematical graphs in the tests [16]. Student solutions were compared with reference plots using the system through the extraction of data points from the answer sheets. Methods, such as mathematical morphology and cross-correlation, aid in obtaining precise evaluations, particularly for linear functions. Although efficient, there are still difficulties in grading the intricate and erratic layouts. This automated method provides a scalable substitute for manual grading of particular graphical jobs. The effectiveness of big language models, in contrast to human grading, is the main emphasis of Flodén [17] evaluation of AI-based grading. Results from an analysis of more than 400 university exam answers revealed that AI assigned more midrange values, with 70% of the results aligning within 10% of the human scores. Although they voiced worries about the accuracy of complex or course-specific questions, they acknowledged that the AI grading was reasonable. The results indicate that, although AI grading shows promise, human oversight is necessary to guarantee reliability and fairness. A fair grading procedure for randomized exams, in which students are given a variety of questions from a vast bank, was examined [18]. Variations in question difficulty result in ex-post unfairness when traditional grading is performed using simple averaging. The proposed technique calculates students' scores across question banks using a maximum likelihood estimator. This can greatly enhance accuracy and fairness. Based on this study, grading bias and mistakes can be reduced by a scalable grading method for randomized tests.

Liu et al. [19] showed how large language models are useful. They are used in math tests to mark handwritten short-answer answers. This can be achieved using AI models, such as GPT-4 and OCR. They highlighted the promising results of automated grading with high accuracy. However, human inspection is still necessary for final grading. This human grading system ensures accuracy for complicated and context-dependent responses. This study suggests that hybrid frameworks can be facilitated by improving the grading standards. Currently, webcams and image processing techniques are used in various exams [20]. This suggests an inexpensive automated grading system for multiple-choice questions. Images were taken to assess the responses on the answer sheet system. These images are then grayscaled, and marked answers are identified using thresholding and edge detection. This is one of the techniques in which fewer manual interactions are performed to assess answers. This method also seeks to increase efficiency and accuracy. This method is also effective in educational institutes that do not have expensive OMR technology. This webcam-based approach offers a scalable way to conduct tests in educational institutions. Various plagiarism detection methods have been described [21]. The greedy string tilting algorithm works well for line-to-line one-word code comparisons. However, this method has limitations for large text databases. The Kolmogorov Complexity Algorithm can be used for difficult scenarios. It can identify patterns based on the minimal matching tokens. Fingerprint is a text-based method that uses a combination of hash tables and tokens. It is most effective for large datasets. Thus, it is a perfect choice for an academic database. However, complex-structure-based Abstract Syntax Tree (AST) and semantic-graph algorithms provide strong logical comparisons. This can improve the accuracy of plagiarism detection in complex circumstances.

## PROPOSED APPROACH

The methodology proposed in Figure 1 addresses the gaps in existing literature and AI models. This is achieved by introducing an AI-driven system for exam grading, percentile calculation, and plagiarism detection. Such systems can be created by integrating Machine Learning (ML) models with external APIs, which can be used to parse documents and retrieve data. This will provide a scalable solution that enhances the integrity of academic assessments. The proposed design and implementation are structured into three sequential phases. This phase helps with the overall functioning and system robustness.



**Figure 1.** Proposed workflow.

These three phases help to develop a comprehensive AI-driven system for academic assessment. First phase of building secure backend infrastructure. It can be accessed via web and mobile applications. This phase ensures the safe storage of exam data, user profiles, and results. External APIs are used to parse PDF documents to enable efficient data handling.

The second phase involves modelling the ML model by training, testing, and deploying it. The deployed model helps in the evaluation of plagiarism detection and student responses. The percentile can be calculated by formula implementation. These models help generate detailed reports for the instructor. This study can be used for accurate grading and academic integrity.

In the last phase, the user-friendly interface is integrated with the ML component and backend. The developed automated system is scalable. This system addresses challenges, such as manual grading effectiveness and unreliable plagiarism detection. This enhances the fairness and efficiency of academic assessment.

## CONCLUSION

Examinations state that AI can be used to revolutionize grading and percentile calculations and to detect plagiarism with the help of data-driven and automated systems. Current systems frequently lack integration, scalability, and the capacity to manage large-scale evaluations fairly and effectively despite significant advancements. By providing educators with a single AI-powered platform to expedite assessment activities, promote integrity, and improve educational results, our proposed method fills these gaps.

Further customization to accommodate various educational environments should be investigated in future studies. The system's capacity to respond to intricate, user-specific inquiries was improved by including sophisticated natural language processing models drawn from a database of scholarly publications, guides, and professional opinions. The adaptability and usefulness of the system in educational institutions can be further enhanced by this advancement.

## REFERENCES

1. Elkhatat AM, Elsaid K, Almeer S. Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *Int J Educ Integr.* 2023;19:17. doi:10.1007/s40979-023-00140-5.
2. Kumar R. Faculty members' use of artificial intelligence to grade student papers: A case of implications. *Int J Educ Integr.* 2023;19(1):9. doi:10.1007/s40979-023-00130-7.
3. Tobler S. Smart grading: A generative AI-based tool for knowledge-grounded answer evaluation in educational assessments. *MethodsX.* 2024;12:102531. doi:10.1016/j.mex.2023.102531.
4. Malik AR, Pratiwi Y, Andajani K, Numertayasa IW, Suharti S, Darwis A, et al. Exploring artificial intelligence in academic essay: Higher education student's perspective. *Int J Educ Res Open.* 2023;5:100296. doi:10.1016/j.ijedro.2023.100296.
5. Lanke MM, Porwal MA, Sarolkar MS, Palaskar MG, Kakad PS. AI based student's assignments plagiarism detector. *Int J Res Appl Sci Eng Technol.* 2022;10(12):991-999. doi:10.22214/ijraset.2022.48080.
6. Perkins M, Gezgin UB, Roe J. Reducing plagiarism through academic misconduct education. *Int J Educ Integr.* 2020;16(1):3. doi:10.1007/s40979-020-00052-8.
7. Elkhatat AM, Elsaid K, Almeer S. Some students plagiarism tricks, and tips for effective check. *Int J Educ Integr.* 2021;17:15. doi:10.1007/s40979-021-00082-w.
8. Foltýnek T, Dlabolová D, Anohina-Naumeca A, Razi S, Kravjar J, Kamzola L, et al. Testing of support tools for plagiarism detection. *Int J Educ Technol High Educ.* 2020;17(1):46. doi:10.1186/s41239-020-00192-4.
9. Swiecki Z, Khosravi H, Chen G, Martinez-Maldonado R, Lodge JM, Milligan S, Selwyn N, Gašević D. Assessment in the age of artificial intelligence. *Comput Educ Artif Intell.* 2022;3:100075. doi:10.1016/j.caeai.2022.100075.

10. Chemaya N, Martin D. Perceptions and detection of AI use in manuscript preparation for academic journals. *PLOS One*. 2024;19(7):e0304807. doi:10.1371/journal.pone.0304807.
11. Weegar R, Idestam-Almquist P. Reducing workload in short answer grading using machine learning. *Int J Artif Intell Educ*. 2024;34(2):247–273. doi:10.1007/s40593-022-00322-1.
12. Xie Y, Wu S, Chakravarty S. AI meets AI: Artificial intelligence and academic integrity—A survey on mitigating AI-assisted cheating in computing education. In: *Proceedings of the 24th Annual Conference on Information Technology Education (SIGITE '23)*; 2023. Marietta (GA): Association for Computing Machinery; 2023. p. 79–83. doi:10.1145/3585059.3611449.
13. Yimyam W, Ketcham M. The grading multiple choice tests system via mobile phone using image processing technique. *Int J Emerg Technol Learn*. 2018;13(10):260. doi:10.3991/ijet.v13i10.8341.
14. Pushpa G, Malleesh A, Siva Ranjan Das D. Enhancing e-assessment with image processing for exam evaluation. *Ind Eng J*. 2022;51(10):100–106.
15. Migut G, Wiersma R. Are we consistent? The effects of digitized exams grading. In: *Proceedings of the 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20)*; 2020. Portland (OR): Association for Computing Machinery; 2020. p. 1338. doi:10.1145/3328778.3372630.
16. Cao Y, Porter L, Liao SN, Ord R. Paper or online? A comparison of exam grading techniques. In: *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '19)*; 2019 Jul 15–17; Aberdeen, Scotland, UK. New York (NY): Association for Computing Machinery; 2019. p. 1–6. doi:10.1145/3304221.3319739.
17. Flodén J. Grading exams using large language models: A comparison between human and AI grading of exams in higher education using ChatGPT. *Br Educ Res J*. 2025;51(1):201–224. doi:10.1002/berj.4069.
18. Chen J, Hartline J, Zoeter O. Fair grading algorithms for randomized exams. [Preprint]. 2023. arXiv:2304.06254v1. doi:10.48550/arXiv.2304.06254.
19. Liu T, Chatain J, Kobel-Keller L, Kortemeyer G, Willwacher T, Sachan M. AI-assisted automated short answer grading of handwritten university level mathematics exams. arXiv [Preprint]. 2024 Aug 21:arXiv:2408.11728.
20. Nithin T, Nasim M, Shekhar TR, Gautam OS, Gholap Y. OMR auto grading system. *Int J Innov Sci Eng Technol*. 2015;2(5):522–526.
21. Grocevs A, Prokofjeva N. Modern algorithms to identify plagiarism. In: *Environment. Technology. Resources: Proceedings of the International Scientific and Practical Conference*. 2019 Jun;Vol 2:61–64. doi:10.17770/etr2019vol2.4058.