مجلة ساحل المعرفة للعلوم الإنسانية والتطبيقية
Sahel Almarifah Journal of Humanities and Applied Sciences
تصدر عن الاكاديمية الليبية فرع الساحل الغربي
المجلد الثاني – العدد الأول - 2026 – الصفحات (302-318)

Comparative Analysis of Cosine Similarity, Jaccard Similarity, and Overlap Similarity for Detecting Cheating Patterns in Online Exams

Musa Faneer1, Abdaroof Abdalla Otman2
[1]Libyan Academy for postgraduate Studies, School of Engineering and Applied Sciences
[2]Alzintan University ,Faculty of information Technology
[1]m.faneer@academy.edu.ly , abdaroof.otman@uoz.edu.ly[2]

التحليل المقارن للتشابه المعتمد على جيب التمام (Cosine) والتشابه وفق مؤشر جاكار (Jaccard)وتشابه التراكب (Overlap) للكشف عن أنماط الغش في الاختبارات الإلكترونية

[1]موسى فنير، عبد الرؤف عبد الله عثمان[2]
[1]الاكاديمية الليبية للدراسات العليا، مدرسة العلوم التطبيقية والهندسية
[2]جامعة الزنتان، كلية تقنية المعلومات

الملخص

تقيّم هذه الدراسة فعالية ثلاث خوارزميات للتشابه: تشابه جيب التمام(Cosine Similarity) ، وتشابه جاكارد (Jaccard Similarity)، وتشابه التداخل(Overlap Similarity) ، في الكشف عن أنماط الغش في الامتحانات الإلكترونية القائمة على نظام مودل .(Moodle) ومن خلال تحليل الإجابات المقالية من مجموعة بيانات تضم 45 طالبًا، يقيّم البحث قدرة كل خوارزمية على الكشف عن مستويات التشابه المتفاوتة التي تشير إلى التعاون غير المصرح به. وتشير النتائج إلى أن خوارزمية تشابه جيب التمام تتفوق على الخوارزميات الأخرى، محققةً دقة (accuracy) قدرها 0.75، ودقة تنبؤية (precision) قدرها 0.75، واسترجاع (recall) قدره 1.0، ومقياس F1 (F1-score) قدره 0.857 في الكشف عن الإجابات المتشابهة. بينما تُعد خوارزمية تشابه جاكارد أكثر فعالية في تحديد حالات التشابه المنخفض (أقل من 50%)، في حين توفر خوارزمية تشابه التداخل نهجًا متوازنًا لحالات التشابه المعتدل (50% - 75%). وتسهم هذه الرؤى في تحسين أنظمة الكشف عن الغش في بيئات التعلم الإلكتروني من خلال اختيار الخوارزمية الأكثر ملاءمة بناءً على خصائص البيانات.

الكلمات المفتاحية :الامتحانات الإلكترونية، الكشف عن الغش، تشابه جيب التمام، تشابه جاكارد، تشابه التداخل.

Abstract

This study investigates the effectiveness of three similarity algorithms, namely Cosine Similarity, Jaccard Similarity, and Overlap Similarity, for detecting cheating patterns in Moodle-based online examinations. The study addresses the challenge of identifying unauthorized collaboration in e-learning environments by analyzing essay responses from a dataset of 45 students and evaluating the performance of different similarity-based detection approaches. The results show that Cosine Similarity achieved the highest

overall performance with an accuracy of 0.75, precision of 0.75, recall of 1.0, and an F1-score of 0.857. Jaccard Similarity ac hieved an accuracy of 0.70, demonstrating better performance in identifying low similarity cases, while Overlap Similarity achieved an accuracy of 0.72, providing a balanced detection capability for moderate similarity levels. These findings highlight the importance of selecting appropriate similarity measures according to data characteristics and contribute to improving automated cheating detection systems in Moodle-based e-learning environments.

Keywords: Online exams, cheating detection, Cosine Similarity, Jaccard Similarity, Overlap Similarity.

Introduction

Introduction The rapid adoption of online learning platforms, such as Moodle, has transformed education by providing students with greater accessibility and flexibility. Online examinations have become an essential component of these platforms, enabling remote assessment and continuous evaluation of student performance. However, the shift from traditional classroom-based exams to online assessments has introduced new challenges related to maintaining academic integrity. The possibility of cheating through copying answers, sharing solutions, or unauthorized collaboration can negatively affect the reliability and fairness of assessment outcomes. Traditional methods for preventing academic dishonesty, such as physical supervision and in-person proctoring, are often difficult to apply in online learning environments. Therefore, automated approaches based on data analysis and intelligent techniques have gained increasing attention as effective solutions for detecting suspicious behaviors. Among these approaches, text similarity analysis provides a promising method for identifying potential cheating by measuring the degree of similarity between students' responses. Different similarity algorithms analyze textual relationships using different mathematical approaches, which may lead to variations in detection performance. Cosine Similarity measures the similarity between text representations based on vector orientation, Jaccard Similarity evaluates the overlap between shared terms, and Overlap Similarity focuses on the proportion of common elements between responses. Due to these differences, comparing these algorithms is important to determine which approach is more suitable for detecting cheating patterns in online exams. This paper evaluates the effectiveness of three widely used similarity algorithms: Cosine Similarity, Jaccard Similarity, and Overlap Similarity,

in detecting cheating patterns in Moodle-based online examinations. The study was conducted at the Faculty of Information Technology at Zintan University using a dataset consisting of 45 student responses. The performance of the algorithms was evaluated using accuracy, precision, recall, and F1-score to identify the most effective approach for detecting suspicious similarities among student answers. The findings provide insights into selecting appropriate similarity-based detection mechanisms for different online exam scenarios and contribute to improving the reliability and integrity of e-learning assessment systems.

Related Work

The increasing adoption of online learning platforms has created a growing need for reliable mechanisms to maintain academic integrity during online examinations. Previous studies have investigated different approaches for detecting cheating behaviors, including statistical analysis, data mining, machine learning, and text similarity-based methods.

Duhaim et al. (2021) [1] investigated cheating detection in Moodle-based online examinations by applying statistical methods, similarity measures, and clustering techniques to analyze student activity data. Their study demonstrated the potential of data-driven approaches in identifying suspicious examination behaviors. Similarly,

Ghizlane et al. (2023) [2] proposed a continuous verification framework for online assessments to improve authentication and reduce unauthorized participation during examinations. Catherine C. et al. (2023) [3] explored collusion detection using text mining and event analysis techniques, showing that analyzing response patterns and interaction behaviors can provide useful indicators of academic dishonesty. In addition to behavioral analysis, text similarity algorithms have been widely applied in educational contexts, particularly for plagiarism detection and identifying similarities between student submissions.

Rinjeni et al. (2024) [4] evaluated the performance of Cosine Similarity and Jaccard Similarity algorithms for academic text matching. Their results showed that similarity measures based on vector representations and shared textual features can effectively identify relationships between documents, while the choice of similarity algorithm

significantly affects detection performance. Furthermore, recent studies have investigated artificial intelligence and natural language processing approaches for improving text similarity detection.

A study published in 2023 [5] proposed an AI-based text similarity detection framework using traditional similarity algorithms and advanced NLP techniques. The study compared approaches such as Cosine Similarity, Jaccard Similarity, TF-IDF-based methods, and deep learning-based semantic models, demonstrating that combining similarity measures with intelligent models can improve the accuracy and robustness of similarity detection systems. More recent research has focused on applying deep learning and AI-based methods for automated cheating detection in online examinations.

Hu et al. (2024) [6]. proposed a multi-perspective adaptive cheating detection system using artificial intelligence techniques to analyze examination behaviors and suspicious patterns. The study highlighted the effectiveness of AI-based monitoring systems in improving online assessment security. Additionally, Shih et al. (2024) [7] introduced an AI-assisted framework for detecting cheating networks in online assessments, emphasizing the importance of intelligent analysis for identifying complex collaboration patterns. Although previous studies have demonstrated the effectiveness of AI, machine learning, and similarity-based techniques in detecting academic misconduct, most existing approaches focus on behavioral monitoring, image-based proctoring, or general plagiarism detection. Limited research has directly compared lightweight text similarity algorithms for detecting cheating patterns in essay-based online examinations. Therefore, this study contributes by evaluating and comparing three widely used similarity algorithms: Cosine Similarity, Jaccard Similarity, and Overlap Similarity, using Moodle-based exam responses to identify the most suitable approach for detecting suspicious similarities among students.

Methodology

The study utilized a dataset of 45 students' exam responses.

Dataset Construction and Ground Truth Definition

The dataset used in this study consists of 45 students' essay responses collected from Moodle-based online examinations at the Faculty of Information Technology, Zintan University. The collected data includes students' answers, question information, and submission details. To establish a reference for evaluating the similarity algorithms, the ground truth was defined based on manually reviewed cases by comparing students' responses and identifying whether the similarity was caused by normal answer similarity or potential unauthorized collaboration. This manual verification process was used as a reference label for evaluating the performance of the proposed similarity detection approaches.

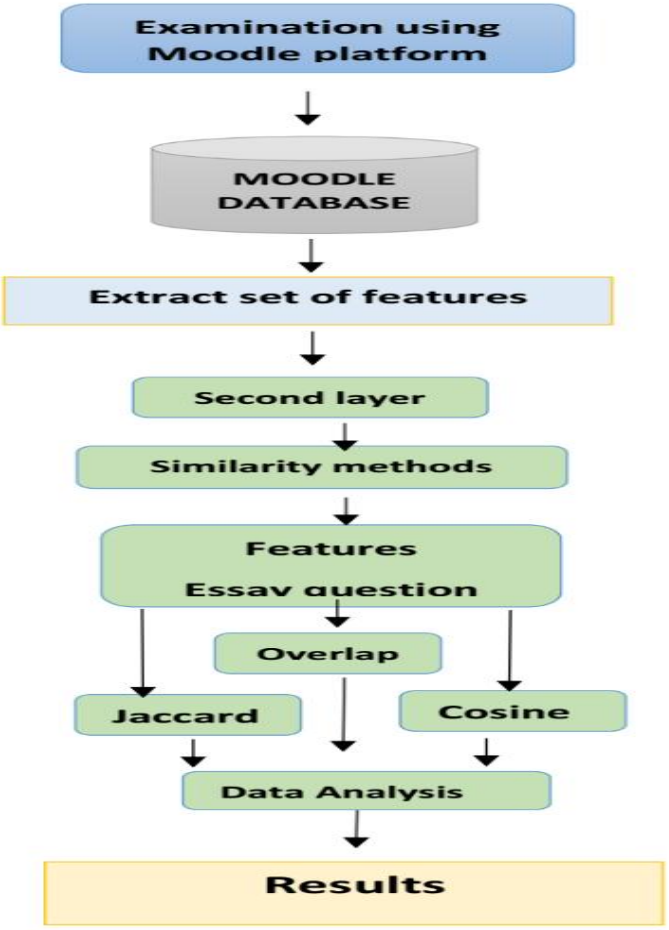


Figure 1:System Methodology

Implementation:

The three similarity algorithms (Cosine Similarity, Jaccard Similarity, and Overlap Similarity) were implemented using Python libraries, including NumPy and scikit-learn.

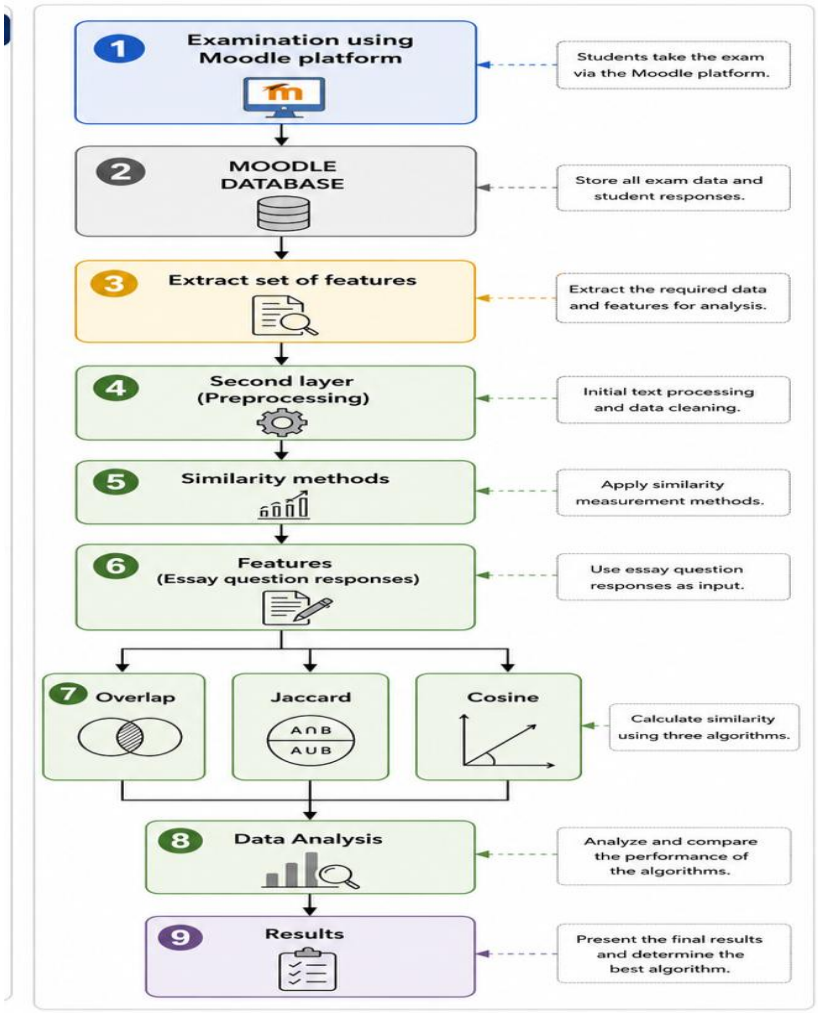


Figure 2 :System Methodology

The proposed methodology focuses on analyzing students' responses in Moodle-based online examinations to detect similarity patterns that may indicate potential cheating or unauthorized collaboration. The process begins by collecting examination data from the

Moodle database, followed by extracting relevant features such as students' answers, student information, and submission details.

The extracted data then undergoes a text preprocessing phase to prepare the responses for analysis by cleaning and organizing the textual data. After that, three similarity algorithms are applied to measure the similarity between students' answers:

1. Cosine Similarity: Measures the similarity between text representations based on vector relationships.
2. Jaccard Similarity: Calculates the proportion of shared terms between two responses.
3. Overlap Similarity: Measures the degree of common elements between student answers.

The similarity scores generated by these algorithms are then analyzed and compared using evaluation metrics such as Accuracy, Precision, Recall, and F1-score to determine the most effective algorithm for detecting suspiciously similar responses.

Finally, the system presents the results and identifies the most suitable similarity algorithm for improving cheating detection in Moodle-based online assessment environments.

Algorithms Used:

Overlap Similarity: Measures the overlap between two sets by dividing the size of their intersection by the size of the smaller set. It balances performance across various text lengths.

$$O(A,B) = \frac{|A \cap B|}{min(|A|,|B|)}$$

Equation 1: Overlap Algorithm Formula based on [1].

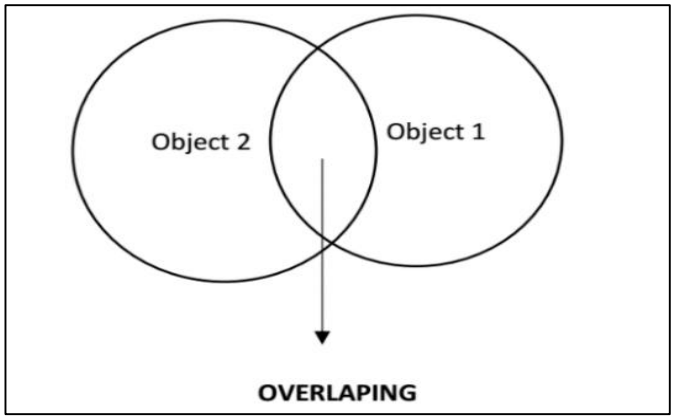


Figure 3: Overlap Algorithm Equation .

Cosine Similarity: Measures the cosine of the angle between two vectors representing text documents. It is effective for long and medium-length texts.

$$C(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}}$$

Equation 2: Cosine Algorithm [1].

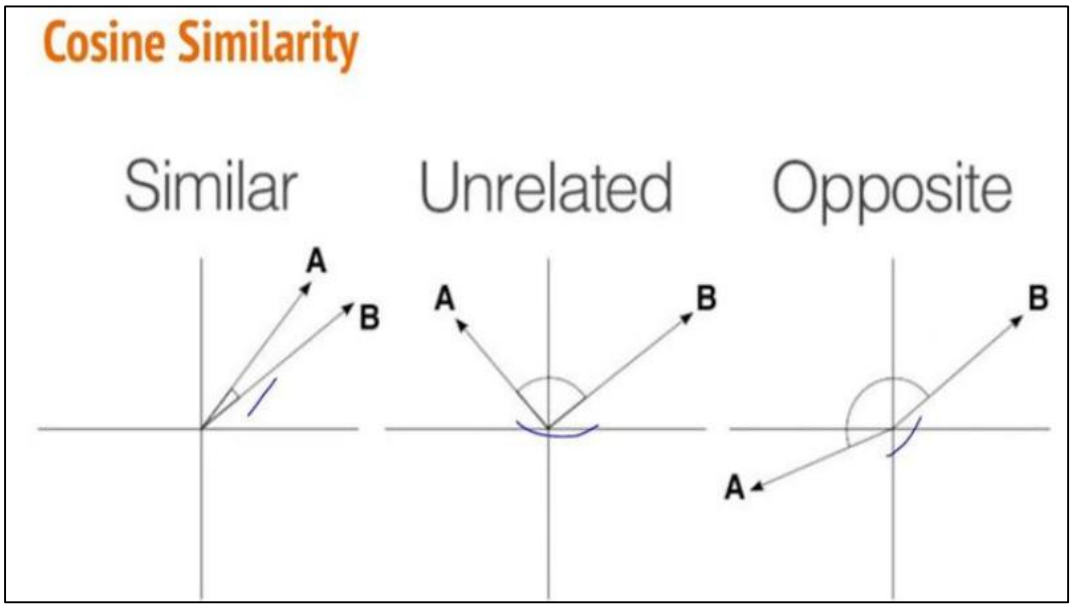


Figure 4: Cosine Similarity [1].

Jaccard Similarity: Calculates the ratio of intersection to union between two sets of words. It is suitable for short texts or keyword-based comparisons.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

Equation 3: Jaccard Algorithm [1].

Design: Analysis of Essay Question Answers

This analysis aims to compare students' answers to the essay question with the model answer, as well as compare the answers among students themselves, using three main algorithms that measure similarity between texts:

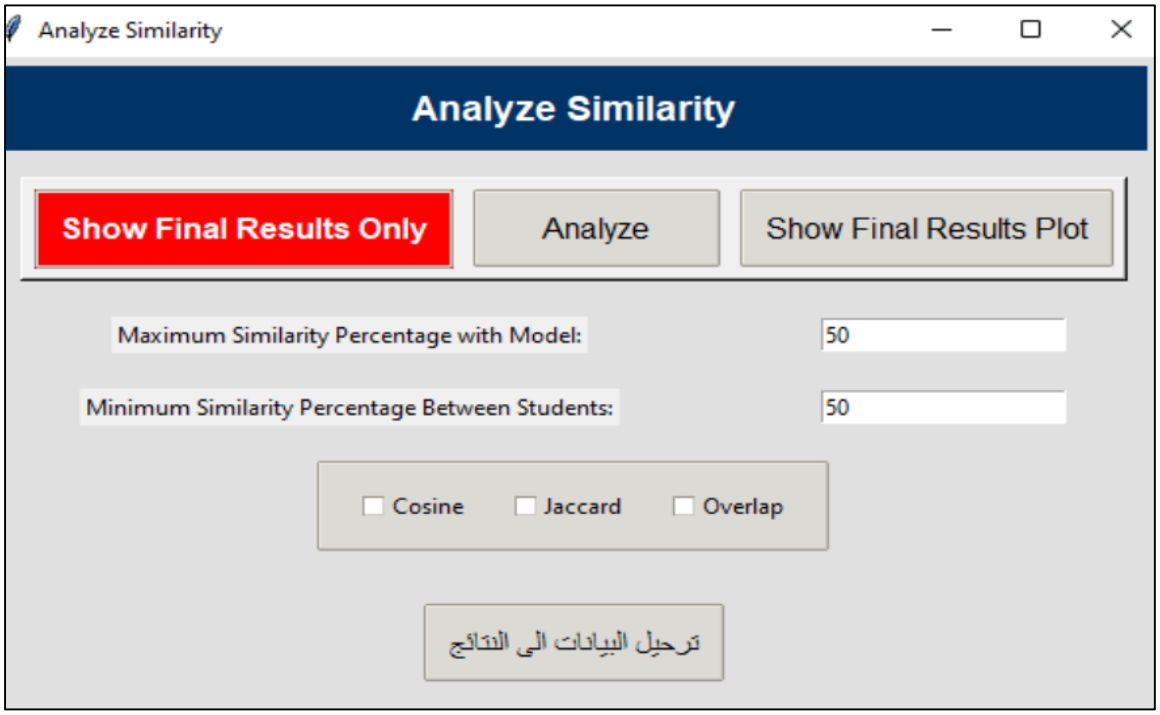


Figure 5: Analysis Similarity.

Application of Algorithms:

By applying these algorithms to the distribution of students' answers, it is possible to accurately detect the level of matching between students' answers and the model answer,

as well as the similarity among students' answers themselves. This helps in identifying cases of cheating, copying, or plagiarism among students.

To ensure effective detection, the following thresholds have been set:

- A maximum similarity threshold of 50% for comparison with the model answer.
- A minimum similarity threshold of 50% for comparison among students' answers.

Based on these thresholds, the results of the three algorithms will be extracted and analyzed to identify potential cases of academic dishonesty.

Similarity Threshold Selection

The similarity thresholds used in this study were selected based on the nature of essay-based examination responses and previous similarity detection practices. A threshold of 50% was considered as an indicator of noticeable similarity between student answers, while 75% represents a high similarity level that may indicate possible copying or unauthorized collaboration. These thresholds help categorize responses into low, moderate, and high similarity levels; however, different examination types and datasets may require adjusting these values to achieve optimal detection performance.

Results:

The results were categorized based on similarity percentages:

High Similarity (>75%):

Cosine Similarity: Detected 8 cases of high similarity.

Jaccard Similarity: Detected 0 cases.

Overlap Coefficient: Detected 13 cases.

Moderate Similarity (50%-75%):

Cosine Similarity: Detected 36 cases.

Jaccard Similarity: Detected 0 cases.

Overlap Coefficient: Detected 31 cases.

Low Similarity (<50%):

Cosine Similarity: Detected 0 cases.

Jaccard Similarity: Detected 42 cases.

Overlap Coefficient: Detected 0 cases.

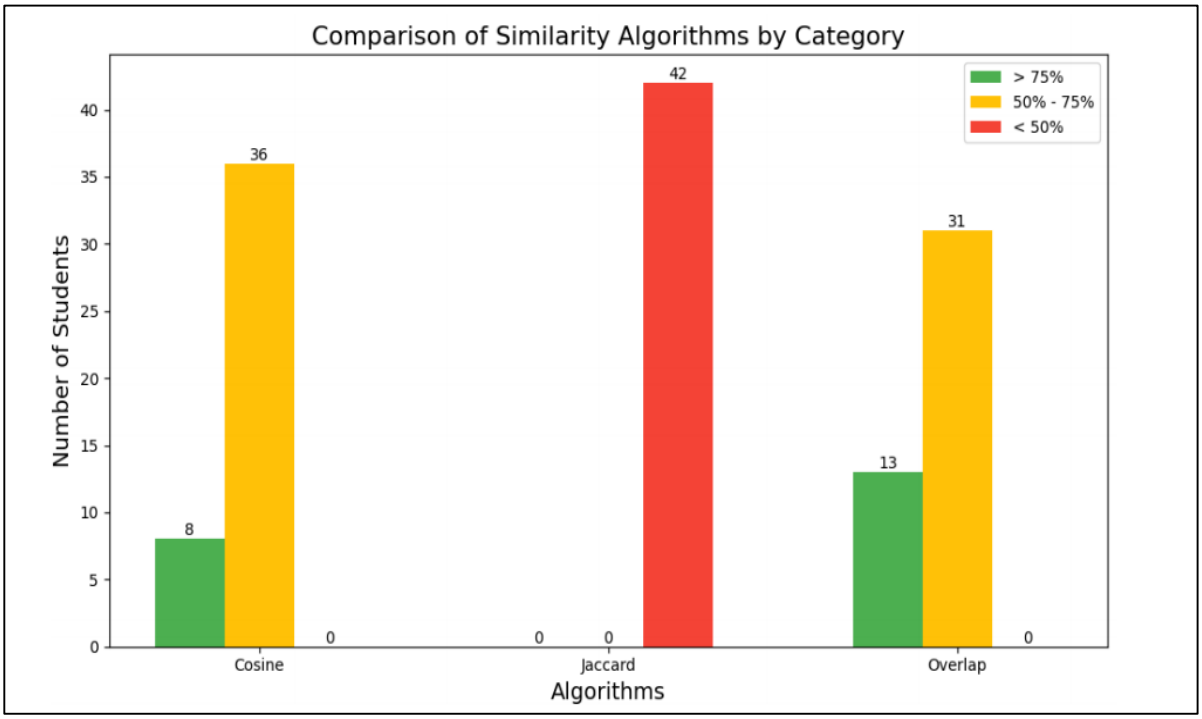


Figure 6: A Comparative Chart of Similarity Algorithms Based on Student Answers.

Evaluation:

A confusion matrix was used to evaluate the accuracy of each algorithm in classifying categories based on similarities:

Set 1: Two completely identical documents.

Set 2: Two documents containing one shared paragraph.

Set 3: Two different documents discussing the same topic.

Set 4: Two completely different documents.

Key performance metrics, including precision, recall, F1-score, and accuracy, were calculated to provide a comprehensive analysis.

Table 1: Similarity Measure Results Between Four Cases

Similarity algorithms	set1	set2	set3	set4
Overlap similarity	100%	54.84%	16.36%	10.29%
Cosine similarity	100%	79.12%	41.57%	29.57%
Jaccard similarity	100%	46.15%	8.16%	6.56%

Interpretation of Similarity Results

Set1 (Two completely identical documents):

All similarity algorithms achieved a similarity score of 100%, which indicates that all methods successfully detected the complete match between identical documents.

Set2 (Documents containing one shared paragraph):

Cosine Similarity achieved the highest similarity score (79.12%), showing better capability in detecting partial similarity in longer texts. Overlap Similarity recorded 54.84%, while Jaccard Similarity achieved 46.15%, indicating lower sensitivity to partial shared content.

Set3 (Different documents discussing the same topic):

Cosine Similarity obtained the highest value (41.57%), demonstrating its ability to capture similarity based on text representation even when the wording is different. Overlap Similarity (16.36%) and Jaccard Similarity (8.16%) produced lower values because they mainly depend on shared words.

Set4 (Completely different documents):

All algorithms produced low similarity scores, confirming their ability to distinguish unrelated documents. Cosine Similarity showed a slightly higher value (29.57%) due to

possible similarity in topic or general terms, while the other algorithms returned lower scores.

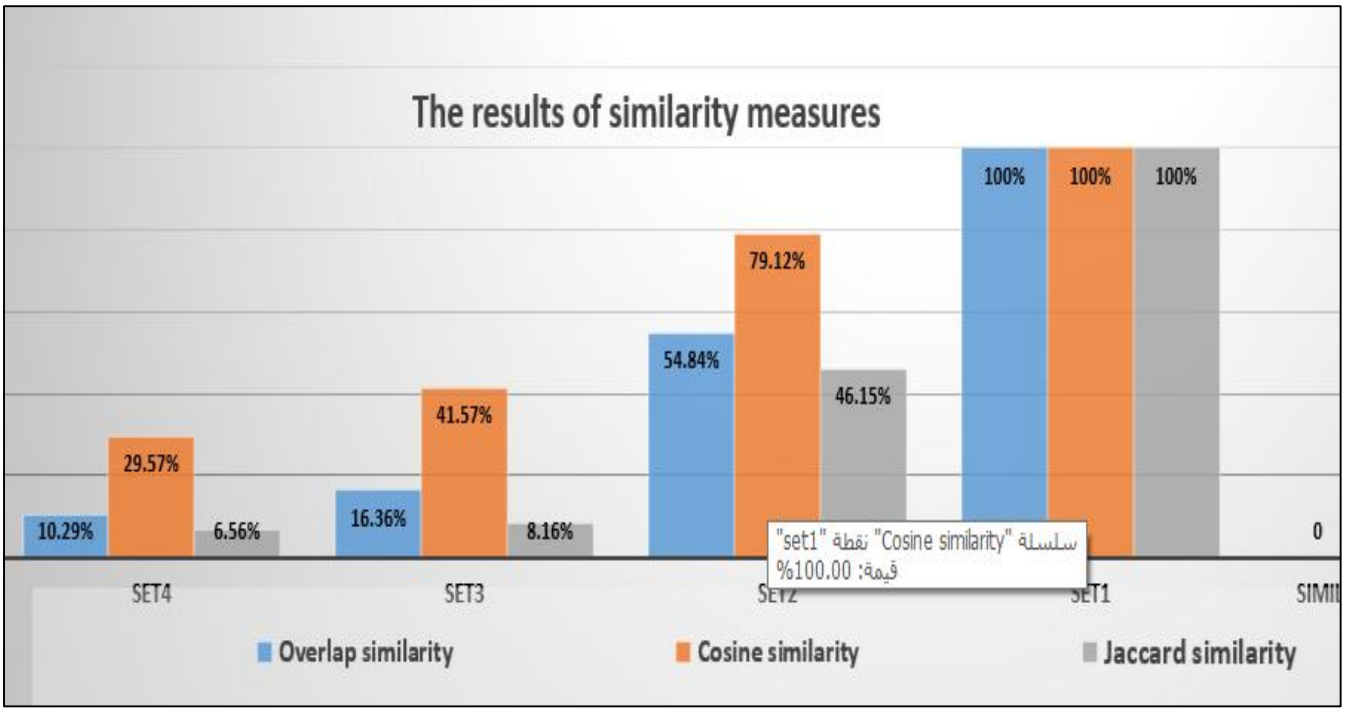


Figure 7: The Chart of Final Comparison of Algorithms.

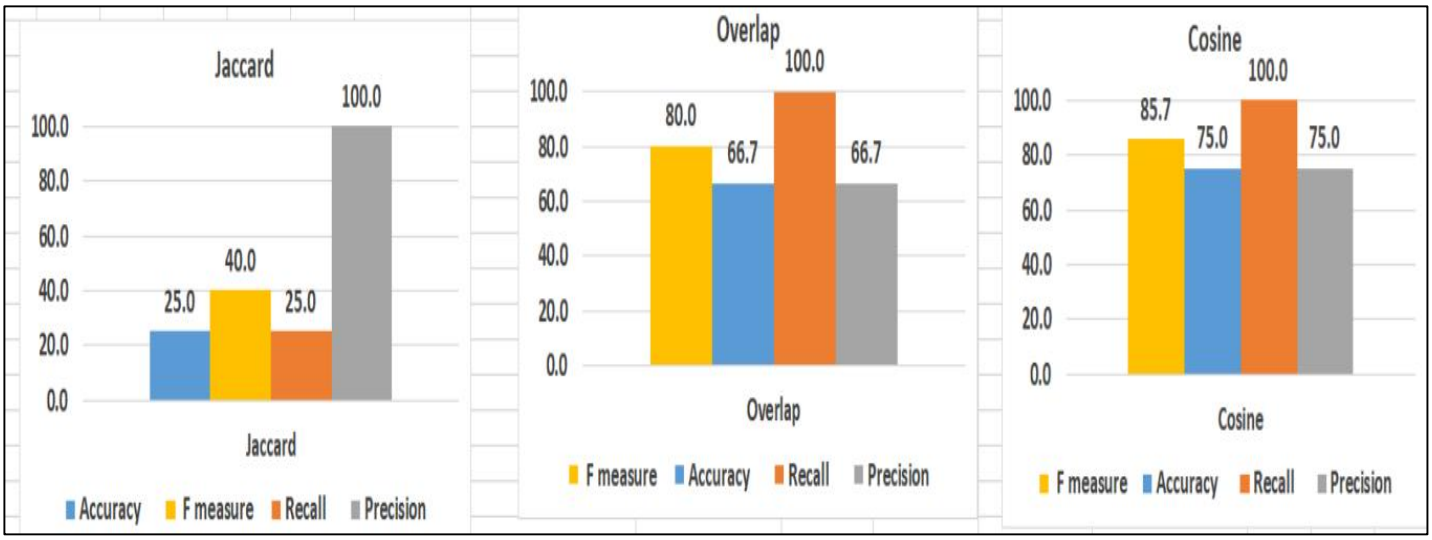


Figure 8: The Chart of Final Comparison of Algorithms.

Discussion

The comparative analysis of similarity algorithms highlights their varying effectiveness in detecting cheating patterns in online exams. Cosine Similarity emerged as the most reliable method, demonstrating superior performance in identifying both high (>75%) and moderate (50%-75%) similarity levels, particularly in essay responses with varying text lengths. In contrast, Jaccard Similarity excelled in detecting low similarity (<50%) but struggled with higher similarity levels, making it more suitable for cases of highly identical documents rather than partial matches. The Overlap Similarity offered a balanced approach, effectively detecting shared paragraphs and moderate similarity levels but proving less effective in cases of very high similarity. Overall, the findings suggest that Cosine Similarity provides the highest accuracy and adaptability for detecting suspicious answer similarities in essay-based online exams. However, the choice of algorithm should align with the nature of the text and the required sensitivity for detecting potential cheating patterns.

Although the study provides valuable insights into the performance of similarity algorithms for cheating detection, the dataset size represents a limitation. The experiments were conducted using responses from 45 students, which may affect the generalization of the findings to larger and more diverse educational environments. Future studies should validate the proposed approach using larger datasets collected from multiple courses, institutions, and examination scenarios to improve the reliability and applicability of the results.

Statistical Analysis

To further support the comparison between similarity algorithms, statistical analysis can be applied to evaluate whether the observed performance differences are significant. Statistical tests such as the paired comparison test can be used to determine whether the performance improvement of one algorithm over others is statistically meaningful rather than resulting from random variation. This provides stronger evidence for selecting the most effective similarity approach for cheating detection.

Key Findings

Performance of Similarity Algorithms:

Cosine Similarity: Identified 8 cases of high similarity (>75%) and 36 cases of moderate similarity (50%-75%). It demonstrated the best performance in analyzing long and medium-length texts.

Precision: 0.75 Recall: 1.0 F1-Score: 0.857 Accuracy: 0.75

Jaccard Similarity: Detected 42 cases of low similarity (<50%) but was less effective in identifying high or moderate similarity levels.

Precision: 1.0 Recall: 0.25 F1-Score: 0.4 Accuracy: 0.25

Overlap Similarity: Identified 13 cases of high similarity (>75%) and 31 cases of moderate similarity (50%-75%). It exhibited a performance level between Cosine and Jaccard Similarity.

Precision: 1.0 Recall: 0.67 F1-Score: 0.8 Accuracy: 0.67

Conclusion

This study presented a comparative analysis of three similarity algorithms, namely Cosine Similarity, Jaccard Similarity, and Overlap Similarity, for detecting potential cheating patterns in Moodle-based online examinations. The experimental results demonstrated that Cosine Similarity achieved the best overall performance, obtaining an accuracy of 0.75, precision of 0.75, recall of 1.0, and an F1-score of 0.857. These results indicate that Cosine Similarity provides a strong balance between detecting suspicious similarities and minimizing missed cases in essay-based responses. Jaccard Similarity achieved an accuracy of 0.25 and showed better capability in identifying low similarity cases, while Overlap Similarity achieved an accuracy of 0.67 and provided a balanced performance for moderate similarity levels. These findings demonstrate that each similarity algorithm has different strengths depending on the characteristics of the text data and the required detection sensitivity.

The results of this study contribute to improving automated cheating detection mechanisms in e-learning environments by providing a comparative evaluation of lightweight similarity-based approaches. The findings can assist educators and system developers in selecting appropriate similarity measures for maintaining academic integrity in online assessments.

Future Works:

Future research should advance online examination cheating detection by integrating advanced machine learning and transformer-based NLP models to enhance the semantic analysis of student responses. To ensure robust generalizability, subsequent studies must expand dataset diversity across varied academic contexts and evaluate algorithmic performance across multiple question typologies. Furthermore, the development of hybrid frameworks that synthesize textual similarity with behavioral metrics such as response times and access patterns is recommended to significantly augment detection accuracy. Additionally, incorporating adaptive similarity thresholds and Explainable AI (XAI) will optimize detection parameters and ensure system transparency for educators. Ultimately, the real-time deployment and empirical evaluation of these methodologies within established learning management systems, such as Moodle, are essential to validate their practical efficacy in large-scale educational environments.

REFERENCES

[1] A. M. Duhaim, S. O. Al-Mamory, and M. Salih, “Cheating detection in online exams during COVID-19 pandemic,” *EasyChair Preprint*, 2021.

[2] G. Moukhliss, R. F. Hilali, and H. Belhadaoui, “Intelligent solution for automatic online exam monitoring,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 5, pp. 5333–5341, Oct. 2023, doi: 10.11591/ijece.v13i5.pp5333-5341.

[3] C. Cleophas, C. Hönnige, F. Meisel, and P. Meyer, “Who’s cheating? Mining patterns of collusion from text and events in online exams,” *INFORMS Transactions on Education*, vol. 23, no. 2, pp. 84–94, 2022, doi: 10.1287/ited.2021.0260.

[4] T. P. Rinjeni, A. Indriawan, and N. A. Rakhmawati, “Matching scientific article titles using Cosine Similarity and Jaccard Similarity algorithm,” *Procedia Computer Science*, vol. 234, pp. 553–560, 2024, doi: 10.1016/j.procs.2024.03.039.

[5] Z. Liu, J. Zhu, X. Cheng, and Q. Lu, “Optimized algorithm design for text similarity detection based on artificial intelligence and natural language processing,” *Procedia Computer Science*, vol. 228, pp. 195–202, 2023, doi: 10.1016/j.procs.2023.11.023.

[6] X. Hu *et al.*, “A multi-perspective adaptive cheating detection system for online examinations based on artificial intelligence,” *Applied Sciences*, vol. 14, no. 10, p. 4048, 2024, doi: 10.3390/app14104048.

[7] M. Shih *et al.*, “Human-in-the-loop AI framework for detecting cheating rings in online assessments,” *arXiv preprint arXiv:2403.14711*, 2024.