Assignment Web Similarity Analysis
Generated on 2025-05-07 14:20:54

Executive Summary

Overall Web Similarity Score: 0%

Assessment: Error performing analysis with Gemini API.

Conclusion: Analysis failed due to an error.

Color Legend for Highlighted Text:

■ High Similarity (75-100%): Potential verbatim copying

■ Medium Similarity (50-74%): Significant paraphrasing

■ Low Similarity (30-49%): Minor similarities

Web Sources Analyzed

Source URL	Similarity Score
https://www.technologyreview.com/2022/12/19/1065596/how-to-spot-ai-generated-text/	41.7%
https://www.eastcentral.edu/free/ai-faculty-resources/detecting-ai-generated-text/	42.22%
https://cloudassess.com/blog/detecting-ai-assessments/	52.96%
https://www.zdnet.com/article/i-tested-10-ai-content-detectors-and-these-5-correctly-identified-ai-text-every-time/	42.05%

Detailed Content Matches

No specific content matches were identified.

Full Assignment with Highlighted Plagiarism

Sections highlighted indicate potential plagiarism, with color indicating similarity level.

Project Title

AI Plagiarism Detector Tool

Group Members

Sewwandi BTI

EG/2020/4210 (Frontend developer)

Balasooriya JM

EG/2021/4424 (UI/UX Designer)

Bandara KMTON

EG/2021/4432 (Mobile APP Developer)

Bandara LRTD

EG/2021/4433 (Backend Developer,Project Manager)

Contact Person

Dr. Kushan Sudheera (kushan@eie.ruh.ac.lk)

Desired Number of Group Members

: 4

Description

The AI Plagiarism Detector Tool aims to develop a system that uses advanced machine learning techniques to identify and prevent plagiarism and detect AI-generated content in student assignments. The tool will serve educational institutions by ensuring the originality and authenticity of submitted work.

Functional Requirements

1.Plagiarism Detection: Check for plagiarism on one or two-page papers. (Expand if feasible) ■ Use departmental assignment submission examples to test. Create algorithms for comparing and evaluating text similarities. ■ Modify any pre-built models as needed. To successfully compare, use techniques such as cosine similarity, word embedding (Word2Vec, GloVe), perplexity, and burstiness.

2.AI-Generated Content Detection: ■ ■ Create techniques to identify AI-generated assignments. ■ Use proper measuring techniques to identify AI-generated text. Use a publicly accessible large language model (LLM) for this purpose if necessary.

3.Submission Portal Development: ■ ■ Develop a user-friendly platform for students to upload documents. ■ Offer fast feedback on plagiarism and AI-generated content scores. Ensure the site displays both results for students following submission.

4.Report Generation: ■■ Create separate papers for AI and plagiarism detections. Ensure documents are downloadable for students and instructors.

5. Database Integration: ■Create a database to hold all scores and related data. Optimize information retrieval and management. This is our project, and our primary functions are assignment to assignment similarity score and assignment to internet similarity score. Please advise me on how to complete that project and provide me with all necessary procedures, as well as this platform similar to Turnity.

Analysis Methodology

Web Similarity Analysis Method: This report analyzes the similarity between a student assignment and web content using multiple approaches:

1. **Basic similarity analysis** using TF-IDF vectorization and cosine similarity metrics to calculate statistical similarity between texts.

2. **Advanced semantic analysis** using Google's Gemini AI to identify conceptual similarities, common phrases, and potential plagiarism patterns.

3. **Source verification** by analyzing multiple sources to distinguish between common knowledge and unique content.

Interpretation Guide:

• 0-15%: Very low similarity - Likely original content

• 16-30%: Low similarity - Contains common phrases but largely original

• 31-50%: Moderate similarity - May contain some paraphrased content

• 51-70%: High similarity - Contains substantial similar content

• 71-100%: Very high similarity - Significant portions may be unoriginal

Disclaimer: This automated similarity analysis provides an approximation of content similarity against web sources. Results should be interpreted by a human reviewer for context-appropriate assessment. Common knowledge, standard phrases, and coincidental matches may be flagged and require human judgment.