Assignment Web Similarity Analysis

Generated on 2025-04-30 15:50:46

Executive Summary

Overall Web Similarity Score: 50%

Assessment: ```json { "overall_similarity_score": 5, "similarity_assessment": "Very low overall similarity. The assignment primarily describes a proposed project, while the web sources are about existing plagiarism detection tools. A few common phrases related to plagiarism detection and software features appear in both, but these are general terms and do not indicate plagiarism.", "detailed_matches": [{ "assignment_text": "Plagiarism Detection", "source_url": "https://www.grammarly.

Conclusion: nt and Format Support", "source_url": "Implied in multiple sources", "source_text": "(Implicit in upload functionality)", "similarity": 20, "match_type": "Common Knowledge" }], "conclusion": "The assignment demonstrates original work. The overlapping phrases are generic terms related to the project's topic and are not considered plagiarism. The student's description of the proposed tool, its features, and their project goals are unique to the assignment." } ```

Web Sources Analyzed

Source URL	Similarity Score
https://www.reddit.com/r/ChatGPT/comments/12syob6/ai_plagiarism_detectors_must_be_stopped_heres_why/	30.77%
https://www.grammarly.com/plagiarism-checker	39.34%
https://gowinston.ai/plagiarism-checker/	46.9%
https://www.plagramme.com/	53.22%

Detailed Content Matches

No specific content matches were identified.

Full Assignment with Highlighted Plagiarism

Sections highlighted in yellow with red text indicate potential plagiarism.

AI Plagiarism Detector Tool under EE5454 software Project Module

Dear Sir,

I hope this message finds you well. We are group of four students eager to take on the “AI Plagiarism Detector Tool” project as part of the EE5454 Software Project Module, with your valuable guidance. Our Team Members are,

EG/2020/4210 - Sewwandi BTI

EG/2021/4424-Balasooriya JM

EG/2021/4432-Bandara KMTON

EG/2021/4433-Bandara LRTD

We envision developing a web application that plagiarism detection with the advanced features.

Advanced test comparison :

Exact matching ,Paraphrase Detection

Content Originality :

Providing the originality of the content as a percentage.

User friendly Interface :

Dashboard, Report Generating

Machine Learning and Natural Language Processing

Using the advanced of the ml algorithms and the NLP techniques to understand the context.

Review and Resubmission

Document and Format Support:

Accesses to upload the various file types (pdf,doc,docx)

User Authentication

We believe this project will greatly benefits to the students as well as the lecturers to their academic works with having high confidence of their academic papers.

Your expertise and feedback would be invaluable as we embark on this journey.

Please let us know your thoughts if we have approval to proceed.

Thank you for your time and consideration.

Best Regards,

Bandara KMTON (EG/2021/4432)

+94 70 428 3880

Department of Computer Engineering,

Faculty of Engineering,

University of Ruhuna

On behalf of the project group members:

Bandara LRTD, Balasooriya JM, Sewwandhi BTI

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.