Document Similarity Analysis Report

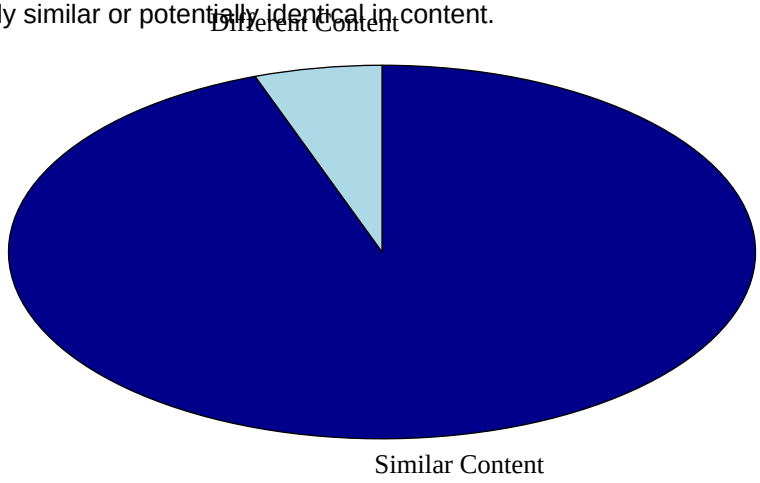
Generated on 2025-03-21 19:39:00

Executive Summary

Overall Similarity Score: 94.48%

Interpretation: The documents are extremely similar or potentially identical in content.

Similarity Visualization



Documents Compared

Property	Document 1	Document 2
Filename	Tharindu_CV_33WX4v7.pdf	Tharindu_Dhanushka_CV_gEF1xOc.pdf
Word Count	687	727

Similar Content Analysis

Top 10 similar sentences/phrases found:

Document 1 Content	Document 2 Content	Match %
Tharindu Dhanushka tharindubandara126@gmail.com +94753923283 github.com	Tharindu Dhanushka tharindubandara126@gmail.com +94753923283 github.com...	100.0%
Proficient in designing intelligent systems, developing scalable web applicat...	Proficient in designing intelligent systems, developing scalable web applicat...	100.0%
Adept at managing end-to-end software development lifecycles and applying inn...	Adept at managing end-to-end software development lifecycles and applying inn...	100.0%
Tools and technologies: Postman, Git, Jira, Agile development	Tools and technologies: Postman, Git, Jira, Agile development	100.0%
Experience Junior Data Scientist (collaborator), Omdena Zambia Chapter - Link...	Experience Junior Data Scientist (collaborator), Omdena Zambia Chapter - Link...	100.0%
May 2024 June 2024 Contributed to Data Collecting Part.	May 2024 June 2024 Contributed to Data Collecting Part.	100.0%
Junior Machine Learning Engineer (Collaborator), Omdena Nigeria Chapter - Lin...	Junior Machine Learning Engineer (Collaborator), Omdena Nigeria Chapter - Lin...	100.0%

Contributed to create Machine Learning Model Using LSTM and Random Forest Reg...	Contributed to create Machine Learning Model Using LSTM and Random Forest Reg...	100.0%
AI Chatbot with Memory Capabilities GitHub Link Developed an intelligent ch...	AI Chatbot with Memory Capabilities GitHub Link Developed an intelligent ch...	100.0%
Implemented multilingual support and deployed a responsive web interface usin...	Implemented multilingual support and deployed a responsive web interface usin...	100.0%

Report Details

Analysis Method: This report uses TF-IDF (Term Frequency-Inverse Document Frequency) vectorization and cosine similarity metrics to analyze document similarity. Additionally, sentence-level comparison is performed using sequence matching algorithms.

Interpretation Guide:

• 0-20%: Very low similarity

• 21-40%: Low similarity

• 41-60%: Moderate similarity

• 61-80%: High similarity

• 81-100%: Very high similarity

Disclaimer: This automated similarity analysis provides an approximation of content similarity. The results should be interpreted by a human reviewer for context-appropriate assessment.