



Multi-Agent Plagiarism-Free Content Rewriter Using Groq LLM and CrewAI

Sahil Mahesh Jamdade

AI & DS

Yashoda Technical Campus, Satara, India

Abstract: Plagiarism has become a growing concern in academic writing and digital content creation due to the easy availability of online resources and the rapid use of artificial intelligence for text generation. While most existing systems focus only on detecting plagiarism, they do not actively help users produce original and meaningful content. This paper presents a Multi-Agent Plagiarism-Free Content Rewriter, an intelligent system designed to rewrite textual content while preserving its original meaning, factual accuracy, and writing intent. The proposed system integrates Groq Large Language Models (LLMs) with a multi-agent framework using CrewAI, where different agents collaboratively perform content extraction, rewriting, and refinement. The system allows users to customize the tone and length of the rewritten content and supports multiple input formats such as text files, PDF documents, and Word files. To assess originality, a lightweight plagiarism estimation method based on shingling and Jaccard similarity is employed.

Index Terms - Plagiarism Reduction, Content Rewriting, Multi-Agent Systems, Groq Large Language Models, CrewAI, Natural Language Processing, Text Similarity Analysis, Academic Writing, AI-Based Content Generation.

I. INSTRUCTION

The increasing availability of digital content has made plagiarism a serious issue in academic writing, research publications, and professional documentation. While plagiarism detection tools are widely used, they mainly focus on identifying similarity rather than helping users create original content. As a result, writers often struggle to maintain originality while preserving the intended meaning and factual accuracy of the source material. To address this challenge, this paper presents a Multi-Agent Plagiarism-Free Content Rewriter that leverages Groq Large Language Models (LLMs) and a CrewAI-based multi-agent architecture. The system assigns distinct roles to multiple agents, including information extraction, content rewriting, and language refinement, enabling controlled and high-quality text transformation. The proposed approach supports multiple input formats, allows customization of tone and length, and provides a plagiarism estimation mechanism using text similarity analysis. Experimental results demonstrate that the system effectively reduces textual similarity while maintaining coherence, readability, and contextual integrity.

II. METHODOLOGY

The proposed system follows a modular approach to generate plagiarism-free rewritten content. Initially, the input text is obtained either through direct user input or by extracting text from uploaded TXT, PDF, or DOCX files. Basic preprocessing is applied to ensure uniform formatting. The rewriting process is performed using Groq Large Language Models (LLMs), optionally enhanced by a CrewAI-based multi-agent framework. In the multi-agent setup, an extractor agent identifies key information, a rewriter agent generates the rewritten text based on user-defined tone and length, and a polisher agent refines the output for clarity and fluency. To estimate originality, the system applies a shingling-based Jaccard similarity method to measure overlap between the original and rewritten texts. The final output is presented to the user and made available for download, ensuring both usability and reliability..

III. FEATURES OF MULTI-AGENT PLAGIARISM-FREE CONTENT REWRITER

The proposed Multi-Agent Plagiarism-Free Content Rewriter provides a comprehensive set of features aimed at generating original, high-quality rewritten content while maintaining usability and reliability. The system supports multiple input methods, allowing users to either paste text directly or upload documents in TXT, PDF, and DOCX formats. Uploaded documents are automatically processed using appropriate text extraction techniques to ensure accurate and consistent input handling.

The core functionality of the system is powered by Groq Large Language Models (LLMs), which perform intelligent content rewriting by preserving the semantic meaning and factual accuracy of the original text. To further enhance rewriting quality, the system optionally employs a CrewAI-based multi-agent framework, where specialized agents collaboratively handle different stages of content transformation. The extractor agent identifies key information and essential terms, the rewriter agent generates a plagiarism-free version of the text based on user-defined tone and length parameters, and the polisher agent improves grammar, coherence, and readability.

The system allows users to customize the rewriting process by selecting the desired tone (such as formal, technical, or casual) and length (short, medium, or long), as well as specifying keywords or phrases that must be preserved exactly. To provide feedback on originality, a lightweight plagiarism estimation module is implemented using text shingling and Jaccard similarity analysis, offering an approximate similarity score between the original and rewritten content.

In addition to content generation, the application features an intuitive web-based interface built using Streamlit, ensuring ease of use for both technical and non-technical users. The rewritten output can be reviewed directly within the application and downloaded in PDF or text format for further use. Overall, these features make the system suitable for academic writing, research documentation, and professional content creation, while ensuring scalability and robustness.

The screenshot displays the user interface of the 'Multi Agent Plagiarism-Free Content Rewriter'. The interface is clean and modern, with a white background and light blue accents. At the top, the title 'Multi Agent Plagiarism-Free Content Rewriter' is prominently displayed, followed by a subtitle: 'Rewrite content to be plagiarism-free while keeping meaning — powered by Groq + CrewAI'. Below this, there are two main sections: '1) Provide Source Content' and '2) Rewriting Options'. The '1) Provide Source Content' section includes a text input field, a file upload area with a 'Browse files' button, and a 'Paste source text here:' text area. The '2) Rewriting Options' section features dropdown menus for 'Tone' (set to 'Neutral') and 'Length' (set to 'Medium'), a text input for 'Preserve exact keywords/phrases (comma-separated)', and a checkbox for 'Ask model to include citations / sources (use '[source needed]' instead of inventing)'. At the bottom, there is a '3) Run' section with two buttons: 'Rewrite Content' and 'Clear Inputs'.

IV. HARDWARE AND SOFTWARE

Hardware Requirements:

Personal computer or laptop
Intel or AMD processor (minimum dual-core)
Minimum 4 GB RAM (8 GB recommended)
At least 10 GB free disk space
Stable internet connection for AI model access

Software Requirements:

Operating System: Windows / Linux / macOS
Programming Language: Python 3.9 or higher
Web Framework: Streamlit
AI Service: Groq Large Language Model (LLM) API
Multi-Agent Framework (Optional): CrewAI
Development Environment: VS Code / PyCharm / any Python IDE

Libraries and Frameworks:

streamlit – Web application interface
groq – Access to Groq LLM API
python-dotenv – Environment variable management
PyPDF2 – PDF text extraction
docx2txt – DOCX document processing
reportlab – PDF generation
re – Text preprocessing and similarity calculation
os and io – File handling operations

V. CONCLUSION

This paper presented a Multi-Agent Plagiarism-Free Content Rewriter designed to address the growing challenge of plagiarism in academic and professional writing. By integrating Groq Large Language Models (LLMs) with a CrewAI-based multi-agent framework, the proposed system effectively rewrites content while preserving semantic meaning, factual accuracy, and writing intent. The use of specialized agents for extraction, rewriting, and refinement enhances the quality and coherence of the generated content. The system supports multiple document formats, offers customizable rewriting options, and provides a lightweight plagiarism estimation mechanism to assess originality. Experimental observations demonstrate that the proposed approach significantly reduces textual similarity while maintaining clarity and contextual relevance. Overall, the system offers a practical, scalable, and efficient solution for generating plagiarism-free content, making it suitable for academic, research, and professional applications..

ACKNOWLEDGMENT

We've worked hard on this project. But without the generous assistance and support of my lecturers and fellow group members, it would not have been feasible. I want to express my gratitude to each and every one of them. Yashoda Technical Campus has my sincere gratitude for their leadership, ongoing oversight, provision of project-related information, and assistance in seeing the project through to completion. I want to thank my parents and the staff at Yashoda Technical Campus for their gracious cooperation and support, which enabled me to finish my project. We have been inspired by their dedication to effective inventory control and customer service. We would like to thank everyone who contributed to the creation of this Multi-Agent Plagiarism-Free Content Rewriter Using Groq LLM and CrewAI once more. This project would not be possible without your help and support.

REFERENCES

- [1] PlagBench: Exploring the Duality of Large Language Models in Plagiarism Generation and Detection Jooyoung Lee¹ Toshini Agrawal² Adaku Uchendu³ Thai Le⁴ Jinghui Chen¹ Dongwon Lee¹
- [2] REWRITELM: AN INSTRUCTION-TUNED LARGE LANGUAGE MODEL FOR TEXT REWRITING Lei Shu[°] Liangchen Luo Jayakumar Hoskere Yun Zhu Yinxiao Liu Simon Tong Jindong Chen Lei Meng[°]
- [3] How Large Language Models are Transforming Machine Paraphrased Plagiarism Jan Philip Wahle * , Terry Ruas* , Frederic Kirstein* , Bela Gipp* *Georg-August-Universität Göttingen, Germany Mercedes-Benz Group AG, Germany
- [4] Paraphrasing evades detectors of AI-generated text, but retrieval is an effective defense Kalpesh Krishna* Yixiao Song Marzena Karpinska John Wieting Mohit Iyyer University of Massachusetts Amherst
- [5] A. Chitra and Anupriya Rajkumar* Plagiarism Detection Using Machine Learning-Based Paraphrase Recognizer
- [6] A Deep Learning Approach for Paragraph-Level Paraphrase Generation for Plagiarism Detection Arwa Al Saqaabi^{1,2} Craig Stewart² · Eleni Akrida² · Alexandra I. Cristea
- [7] Paraphrase Generation: A Survey of the State of the Art Jianing Zhou and Suma Bhat University of Illinois at Urbana-Champaign {zjn1746, spbhat2}@illinois.edu
- [8] Beyond Black Box AI-Generated Plagiarism Detection: From Sentence to Document Level Mujahid Ali Quidwai New York University maq4265@nyu.edu Chunhui Li Columbia University cl4282@columbia.edu Parijat Dube IBM Research pdube@us.ibm.com

