

PLAGIARISM DETECTOR

MINI PROJECT

ARTIFICIAL INTELLIGENCE

TEAM MEMBERS

- ANAHITA GUPTA (RA2011003010189)
- TRAYI KASHYAP (RA2011003010162)

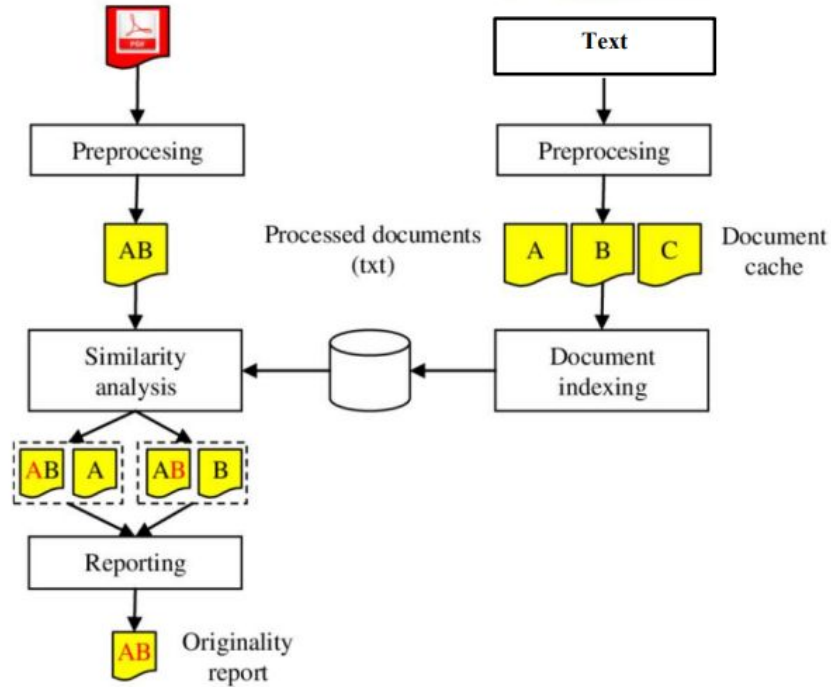
ABSTRACT

Plagiarism is a significant problem in academic and professional writing, and plagiarism detectors play a crucial role in detecting instances of plagiarism. Plagiarism detectors use various algorithms and techniques to compare submitted work against a large database of existing sources to identify any instances of plagiarism. These tools have become increasingly popular among educators and institutions due to the rise of online learning and the ease of access to digital information. There are several types of plagiarism detectors available, ranging from basic free tools to advanced paid versions that provide more detailed analysis. Basic plagiarism detectors compare submitted work against a limited number of sources, while more advanced tools can compare against a vast database of sources and provide a detailed analysis of potential plagiarism. In conclusion, plagiarism detection tools play a crucial role in ensuring academic integrity and preventing plagiarism in various fields. These tools have become increasingly popular among educators and institutions and are continually evolving to provide better accuracy, deeper analysis, and integration with learning management systems.

INTRODUCTION

Plagiarism detector is a tool designed to identify instances of plagiarism in written work. Plagiarism, the act of using someone else's work without giving proper credit or attribution, is a serious offense in academic and professional writing that can lead to severe consequences, including legal action, loss of reputation, and academic failure. Plagiarism detection tools use various algorithms and techniques to compare submitted work against a large database of existing sources to identify any instances of plagiarism. First and foremost, they promote academic integrity by identifying instances of plagiarism and ensuring that students understand the importance of proper citation and attribution. While plagiarism detection tools are not foolproof, they can provide valuable assistance to educators in detecting and addressing instances of plagiarism. There is always the possibility of false positives or false negatives, but these tools remain an essential resource for maintaining academic standards and upholding the value of original work. Overall, plagiarism detector is a crucial tool in promoting academic integrity and preventing plagiarism. It is an essential resource for educators and institutions, providing a way to ensure that all work is original and properly cited.

SYSTEM ARCHITECTURE



METHODOLOGY

1 Project Objective

The objective of the project is to develop a web application for plagiarism detection. Given an input text, the application will search the web for similar content, calculate similarity scores, and generate a report highlighting potential instances of plagiarism.

2 Web Application Setup

Set up a Flask web application to provide an interface for users to input text and view the plagiarism detection report.

3 Index Page

Create an index page where users can enter the text to be checked for plagiarism. Render an HTML form to accept the input text.

4 Text Processing and Similarity Calculation

Utilize the functions in ``similarity.py`` to process the input text and calculate similarity scores against web content. The ``purifyText()`` function removes stop words from the input text, and ``webVerify()`` performs web searches to retrieve relevant URLs. The ``similarity()`` function calculates the similarity between the input text and the extracted web content.

5 Generating the Plagiarism Report

Adapt the `report()` function in `similarity.py` to generate a plagiarism detection report. Iterate through the matching websites and calculate similarity scores for each. Sort the matches based on similarity scores in descending order. Include additional information, such as source URLs and percentage of similarity, to provide comprehensive details about potential instances of plagiarism.

6 Result Page

Create a result page that displays the plagiarism detection report. Render the report in the form of an HTML table generated by the `returnTable()` function in `similarity.py`. Present the matching websites, their similarity scores, and other relevant information to the user.

7 Web Search and Text Extraction

Utilize the functions in `websearch.py` to perform web searches and extract text content from the web pages. The `searchBing()` function searches Bing using the input text and retrieves a specified number of relevant URLs. The `extractText()` function retrieves the HTML content of a given URL and extracts the text content.

8 Flask Application Execution

Run the Flask application, allowing users to access the plagiarism detection functionality through the web interface. Enable debugging mode for easy troubleshooting during development.

SCREENSHOTS AND RESULTS

Plagiarism Detector

Paste your text here:

Rock-cut sculptures are a form of sculpture created by carving directly into natural rock formations, usually on cliff faces, cave walls, or other rock surfaces. This technique has been used by various civilizations throughout history, including in ancient India, Greece, and Egypt. In contrast, stone sculptures are created by carving, chiseling or shaping stones, which have been quarried from rock formations

Generate Report

	Similarity (%)
https://en.wikipedia.org/wiki/Indian_rock-cut_architecture	0.381608
https://whc.unesco.org/en/tentativelists/6628/	0.480769
https://www.britannica.com/topic/Indus-civilization	0.766333
https://www.stoneworld.com/articles/85376-exploring-the-stone-carving-traditions-of-india	1.277839

CONCLUSION

In conclusion, a plagiarism detector is a crucial tool for ensuring academic integrity and preventing plagiarism. The use of plagiarism detection software has become increasingly popular among educators and institutions due to the rise of online learning and the ease of access to digital information. Plagiarism detection tools use various algorithms and techniques to compare submitted work against a large database of existing sources to identify any instances of plagiarism. While these tools are not foolproof, they can provide valuable insights into potential cases of plagiarism that might otherwise go unnoticed.

FUTURE ENHANCEMENTS

As technology continues to advance, there is potential for further enhancements to plagiarism detection software. Here are some possible areas for improvement:

- Improved accuracy: Future enhancements could focus on improving the accuracy of the algorithms used, which could result in more reliable and accurate detection of plagiarism.
- Multilingual detection: Future enhancements could focus on developing multilingual plagiarism detection tools that can identify instances of plagiarism in a variety of languages.
- Integration with learning management systems: Future enhancements could focus on integrating plagiarism detection tools with popular LMS platforms, making it easier for educators to use the tools and track student progress.
- Better student education: While plagiarism detection tools can help identify instances of plagiarism, it's also important to educate students on the importance of academic integrity and proper citation practices.

THANK YOU