PLAGEARLY -- A Plagiarism Detection Platform

A robust system that accurately identifies plagiarism, generates detailed similarity reports, and detects paraphrasing.

OVERVIEW

- The increasing importance of originality in research demands effective plagiarism detection tools. Our platform provides a comprehensive solution tailored for academic integrity

- Many tools are either too complex or costly. This platform offers a streamlined and accessible alternative.

CORE FEATURES

1. **NLP-Based Analysis**: Uses basic natural language processing techniques to analyze text similarity.

2. **Similarity Reports**: Highlights matched sections and provides similarity percentages.

3. **Paraphrase Detection**: Basic rephrasing detection using synonym and sentence structure analysis.

4. **User Data Storage**: Allows users to log in and save their plagiarism reports and data in the database for future reference.

TECHNOLOGY

- ***Natural Language Processing (NLP):*** Identifies word patterns and detects close rephrasing without complex algorithms.

1. **NLTK (Natural Language Toolkit)**: A fundamental library in Python for text processing, tokenization, stemming, and synonym matching, essential for identifying paraphrased content and analysing sentence structures.

2. **Scikit-Learn**: A machine learning library that provides tools for basic similarity detection algorithms, vectorization, and classification, making it easier to implement efficient text comparison models.

- ***Database Search:*** Compares the input document with a set of stored references for matching content.

BENEFITS

- ***Simple and Effective***: Provides essential detection without advanced complexities.

- ***User-Friendly Reports***: Easy-to-read similarity and paraphrase indicators.

- ***Cost-Effective:*** A basic tool requiring minimal resources.

LIMITATIONS & FUTURE ENHANCEMENTS

Limitations:

1. Limited to basic NLP; may miss advanced paraphrasing or complex matches.

Future Enhancements:

1. Potential to integrate more advanced ML algorithms and larger databases as the platform grows.
2. **Pricing System**: Introducing a subscription model with:

   - **Basic Plan**: Offers essential plagiarism detection but with limited features.
   - **Pro Plan**: Provides advanced tools, detailed reports, and additional analysis options for users who need in detailed comprehensive checks.