



Department of Information Technology

NBA Accredited

A.P. Shah Institute of Technology

— G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic Year 2021-2022

A Project Report on
Plagiarism Detection System

Submitted in partial fulfilment of the degree of
Bachelor of Engineering(Sem-6) in

INFORMATION TECHNOLOGY

By

Aditya Joshi(19104044)

AbhayPratap Singh(19104037)

Bharat Singh(19104043)

Under the Guidance of
Prof. Rajashri Chaudhari

1. Project Conception and Initiation

1.1 Objectives

- To avoid plagiarism by ensuring the proper citing of information from outside sources.
- To help users know how to use sources ethically and effectively through summarizing, paraphrasing and quoting.
- To help users understand how to make use of information from outside sources while maintaining a strong voice of their own.

1.2 Literature Review

Hayden Cheers, Yuqing Lin and Shamus P. Smith	Academic Source Code Plagiarism Detection by Measuring Program Behavioral Similarity - 29 th March, 2021	Approaches to measuring the similarity of two programs can be broadly classified by what aspects of the programs are being compared
John L. Donaldson, Ann-Marie Lancaster and Paula H. Sposato	Plagiarism Detection System -	The problem of students handing in research papers/journals which are not their own has caused concern among both faculty and students.

1.3 Problem Definition

- Plagiarism is the representation of someone else's work as one's own and it has become a very serious problem in academics now-a-days.
- Manual comparison is rather impractical as well as difficult considering the amount of work that would be needed to be done.
- One way of avoiding plagiarism is by citing the correct external source from where the information was taken.

1.4 Scope

- This application can be used to read author manuscripts and student papers in a few minutes, matching what they have submitted to already published work.
- The application highlights the parts that match/are similar to already existing files/documents.
- The text that is highlighted includes content that is identified as too similar to published content. The highlighted text is used to generate a score that represents the percentage of similarity between the author's document and published content.

1.5 Technology stack

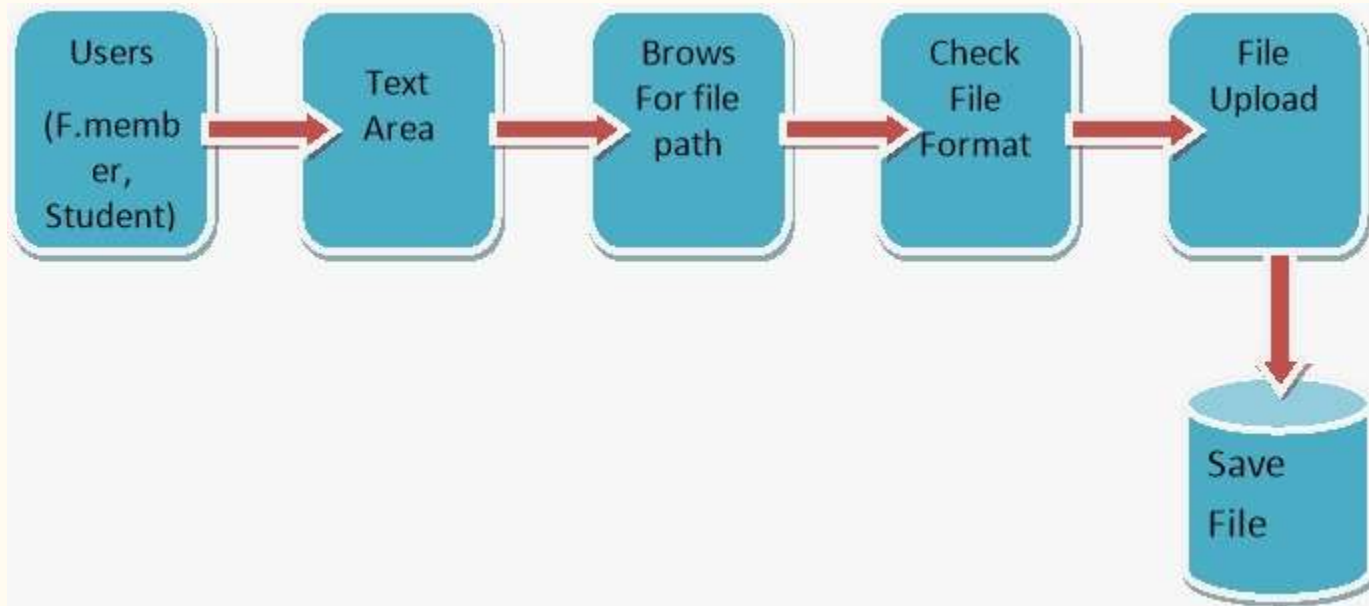
- Python
- Tkinter
- SequenceMatcher Algorithm

2. Project Design

2.1 Proposed System

- Using our application users can either upload the files to check for plagiarism. After that application will check for file format because the uploaded files will be in txt format only.
- If the file format is acceptable then it will show that how much content is plagiarised with respect to the dataset.
- Users can also directly put the content in the text area for checking the plagiarism.

2.2 Design(Flow Of Modules)



3. Implementation

Plagiarism Detector

Insert your text here

0

Submit

Reset

Plagiarism Detector

Insert your text here

Plagiarism is the representation of another author's language, thoughts, ideas, or expressions as one's own original work.[1][2] In educational contexts, there are differing definitions of plagiarism depending on the institution.[3] Plagiarism is considered a violation of academic integrity and a breach of journalistic ethics. It is subject to sanctions such as penalties, suspension, expulsion from school [4] or work,[5] substantial fines[6][7] and even imprisonment.[8][9]

Generally, plagiarism is not in itself a crime, but like counterfeiting, fraud can be punished in a court [10][11] for prejudices caused by copyright infringement,[12][13] violation of moral rights,[14] or torts. In academia and industry, it is a serious ethical offense.[15][16] Plagiarism and copyright infringement overlap to a considerable extent, but they are not equivalent concepts,[17] and many types of plagiarism do not constitute copyright infringement, which is defined by copyright law and may be adjudicated by courts.

Submit

Reset

Activate Windows

Go to Settings to activate Windows.

5. Result

Plagiarism Detector

Insert your text here

Plagiarism is the representation of another author's language, thoughts, ideas, or expressions as one's own original work.[1][2] In educational contexts, there are differing definitions of plagiarism depending on the institution.[3] Plagiarism is considered a violation of academic integrity and a breach of journalistic ethics. It is subject to sanctions such as penalties, suspension, expulsion from school [4] or work,[5] substantial fines[6][7] and even imprisonment.[8][9]

Generally, plagiarism is not in itself a crime, but like counterfeiting, fraud can be punished in a court [10][11] for prejudices caused by copyright infringement,[12][13] violation of moral rights,[14] or torts. In academia and industry, it is a serious ethical offense.[15][16] Plagiarism and copyright infringement overlap to a considerable extent, but they are not equivalent concepts,[17] and many types of plagiarism do not constitute copyright infringement, which is defined by copyright law and may be adjudicated by courts.]

Submit

Reset

test1.txt - 2.29

save.txt - 1.6

records.txt - 1.57

1000.txt - 0.92

userdet.txt - 0.92

file.txt - 0.57

Activate Windows

Go to Settings to activate Windows.

6. Conclusion and Future Scope

User can see the Plagiarised Content using our Application.

User also has options to check their content either by typing in the text area or by uploading the file directly.

Percentage is also showed with respect to the file for plagiarised content.

References

- https://www.tutorialspoint.com/python/python_gui_programming.html
- <https://github.com/search?q=plagiarism+detector>
- <https://www.educative.io/edpresso/what-is-sequencematcher-in-python>
- Academic Source Code Plagiarism Detection by Measuring Program Behavioral Similarity
- Plagiarism Detection System

Thank You

