

Plagiarism Scan Report

Report Generated on: Dec 20,2022

<div><div>0%</div><div>Plagiarised</div></div>	<div><div>100%</div><div>Unique</div></div>	<div><div>Total Words:269</div><div>Total Characters:1699</div><div>Plagiarized Sentences:0</div><div>Unique Sentences:14 (100%)</div></div>
--	---	--

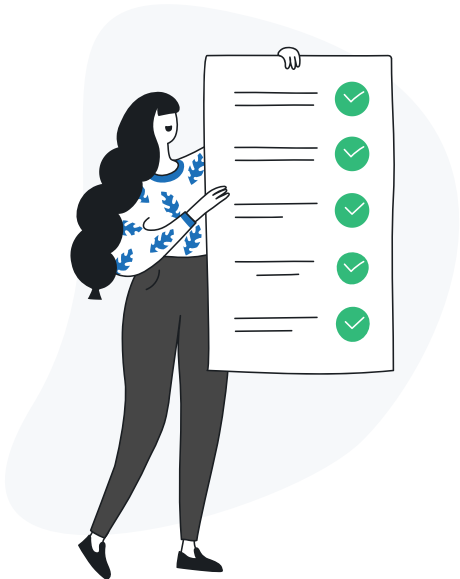
Content Checked for Plagiarism

In this module, the news dataset from kaggle which contains attributes (Title, Text, Label (REAL/FAKE)) are taken. Null value records are eliminated during preprocessing work. Unicode characters are removed from the paragraphs. The dataset download link (https://drive.google.com/file/d/1XoHoUjZtLQ6gh_CPv9_igtZoEsdkgq3t/view?usp=sharing).

In this module, a string comparison is being used which checked the whole sentence of the given news is checked against with all the records in the dataset paragraphs. Removal of Unicode characters are not carried out. So if the full news content is present in the dataset records, then it will be treated as fake or duplicate news information. Whole string comparison is carried out and so partial percentage matching is not applied here. So processing time will be more, if the dataset contains more records.

In this module, a string comparison is being used which checked the sentence of the given news is checked against with all the records in the dataset paragraphs. Removal of Unicode characters are carried out. Even if the partial news content is present in the dataset records, then it will be treated as fake or duplicate news information with percentage match. Partial string comparison is carried out and so percentage matching percentage can be found out here.

In this module, 80% of the data in given data set is taken as training data and 20% of the data is taken as test data. The text (categorical) columns are converted into numerical values. Then the model is trained with training data and then predicted with test data. Of which, most of the records are classified as real or fake and fewer records are classified as Suspicious.



No Plagiarism Found

