**MONISH R**

**2021503030**

# NAAN MUDHALVAN PROJECT REPORT

**TITLE – NEWSPAPER ARTICLE SUMMARY GENERATOR**

# PROBLEM STATEMENT:

In today's fast-paced world, keeping up with the latest news from diverse sources can be overwhelming. This project presents a Newspaper Summary Generator, an AI-powered tool designed to aid readers in quickly digesting key information from news articles. The generator utilizes natural language processing (NLP) techniques to extract essential content, condensing lengthy articles into concise summaries.

The system begins by fetching news articles from various sources, leveraging the newspaper3k library for article scraping. Once the articles are retrieved, the system employs NLP algorithms to analyze and summarize the content. Techniques such as sentence tokenization, keyword extraction, and sentence ranking are utilized to identify crucial information.

To generate the summaries, the system ranks sentences based on their relevance and importance within the context of the article. The sentences with the highest significance are then stitched together to form a coherent and informative summary. Users have the flexibility to specify the number of sentences or paragraphs they wish to include in the summary, tailoring the output to their preferences.

The Newspaper Summary Generator aims to enhance the efficiency of news consumption, allowing users to stay informed without the need to sift through extensive articles. By providing condensed yet comprehensive summaries, this tool empowers readers to grasp the essence of news stories quickly and efficiently.

# OBJECTIVES:

* To facilitate efficient news consumption. The project aims to address this challenge by providing users with a tool that condenses lengthy news articles into concise and informative summaries.
* To save users valuable time by eliminating the need to read through entire articles to extract key information. With the Newspaper Summary Generator, users can quickly obtain an overview of news stories, enabling them to stay informed without investing significant time.
* To make news content more accessible to a wider audience. The generator caters to users who may have limited time to read lengthy articles or those who prefer to receive summarized information. This inclusivity allows individuals with diverse schedules and preferences to access and engage with news content.

**LIBRARY MODULES:**

* + nltk: Provides tools for working with human language data (text).
  + newspaper: A library for downloading and parsing articles.
  + sent\_tokenize: Tokenizes text into sentences.
  + stopwords: A list of common words to ignore in text processing.
  + cosine\_distance: Measures the cosine of the angle between two vectors, used for similarity.
  + numpy: Provides support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays.
  + networkx: A Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks.

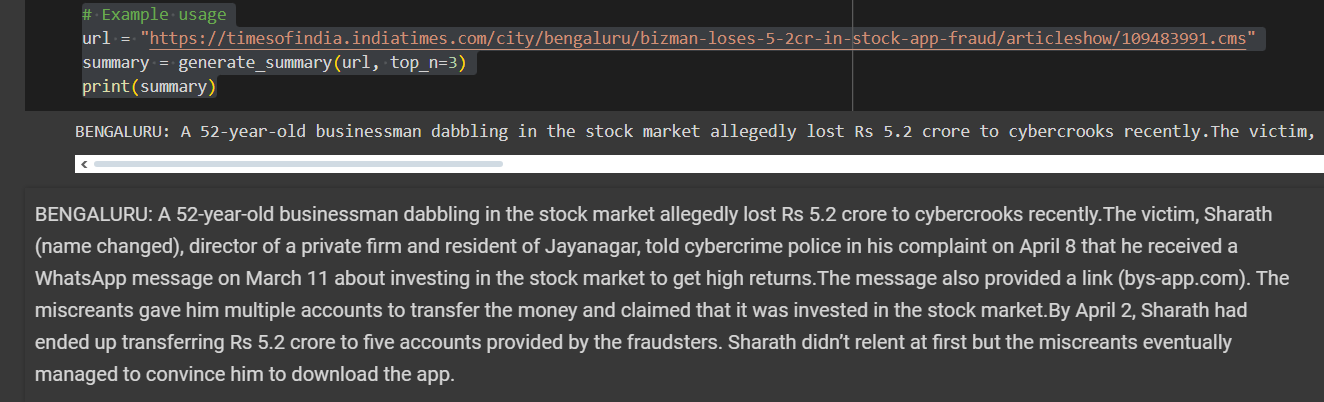
**OUTPUT:**

# Example usage

url = "https://timesofindia.indiatimes.com/city/bengaluru/bizman-loses-5-2cr-in-stock-app-fraud/articleshow/109483991.cms"

summary = generate\_summary(url, top\_n=3)

print(summary)

****