

CIA3-NLP

NAME: DHARAMRAJ BHATT
REG.NO: 1947216

1) TITLE

Text Summarization

2) AIM

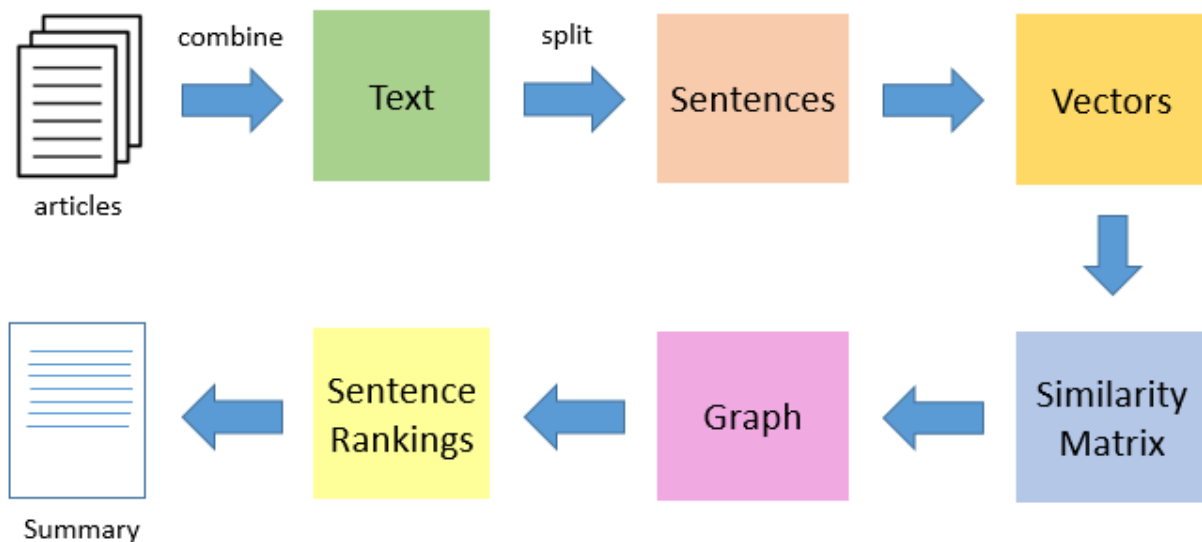
The main of this is to summarize the different articles on tennis using the TextRank Algorithm

3) OBJECTIVE

The objective is to write a python code to do the work as the famous app named Inshort is doing. Inshort is an innovative news app that converts news articles into a 60-word summary.

4)METHODOLOGY/PROCEDURE/ OF YOUR CASE STUDY/APPLICATION BASED PROGRAM(CIA3)

TextRank Algorithm: TextRank is an extractive and unsupervised text summarization technique. Let's take a look at the flow of the TextRank algorithm that we will be following:



- The initial step would be to concatenate all of the text in the articles.
- After that, the text would be divided into individual sentences.

- We'll find vector representations (word embeddings) for each sentence in the next phase.
- Following that, the similarity between sentence vectors is determined and stored in a matrix.
- For sentence rank calculation, the similarity matrix is turned into a graph with sentences as vertices and similarity scores as edges.
- Finally, the final summary is made up of a set of top-ranked sentences.