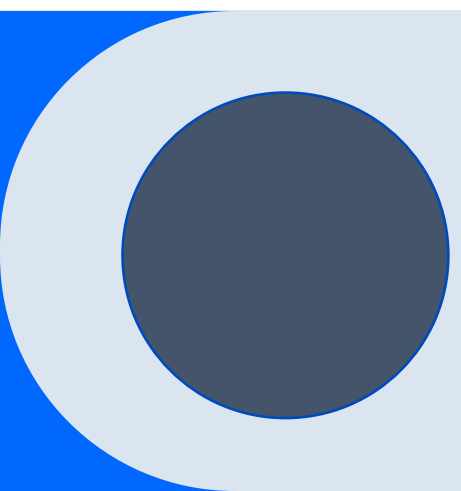




Text Summarization Using NLP



Presented By.

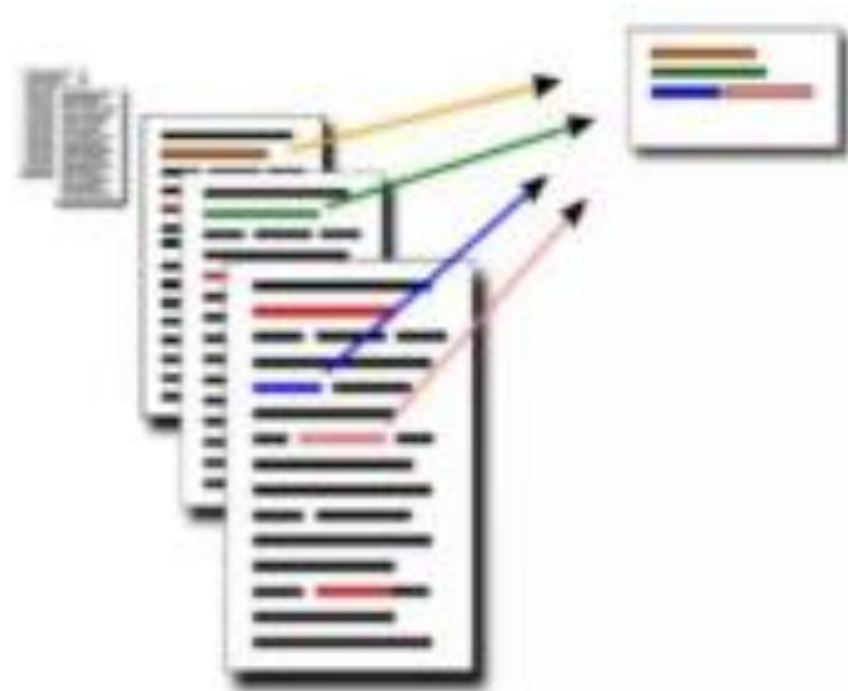
Abhishek Singh



Introduction

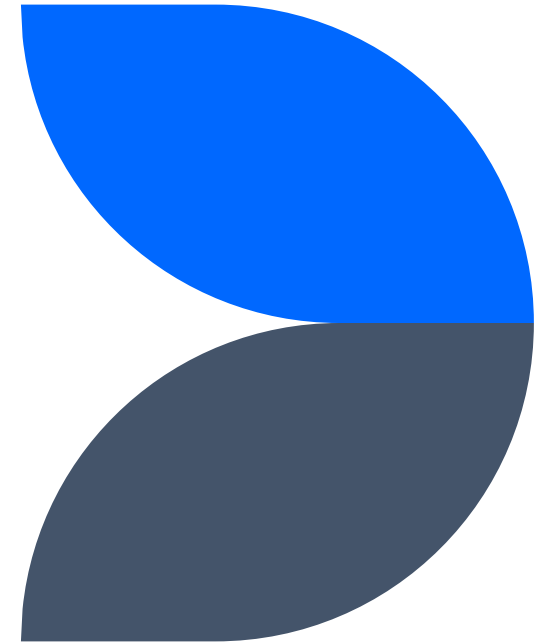
The goal of summarization is to produce a shorter version of a source text by preserving the meaning and the key content of the original document.

A well written summary can significantly reduce the amount of work needed to digest large amount of text.



Types of Text Summarization

- Extractive Summaries
- Abstractive Summaries



Extractive Summaries

- Extractive summaries are created by reusing portions (words, sentences, etc.) of the input text document
- The System extracts text from the entire collection, without modifying the text document,
- Most of the summarization research today is on extractive summarization.

Abstractive Summaries

- Requires deep understanding and reasoning the text.
- It provides own summary over input text without using same word or sentence in the text.
- Determines the actual and short meaning of each element, such as word, sentences and paragraphs.

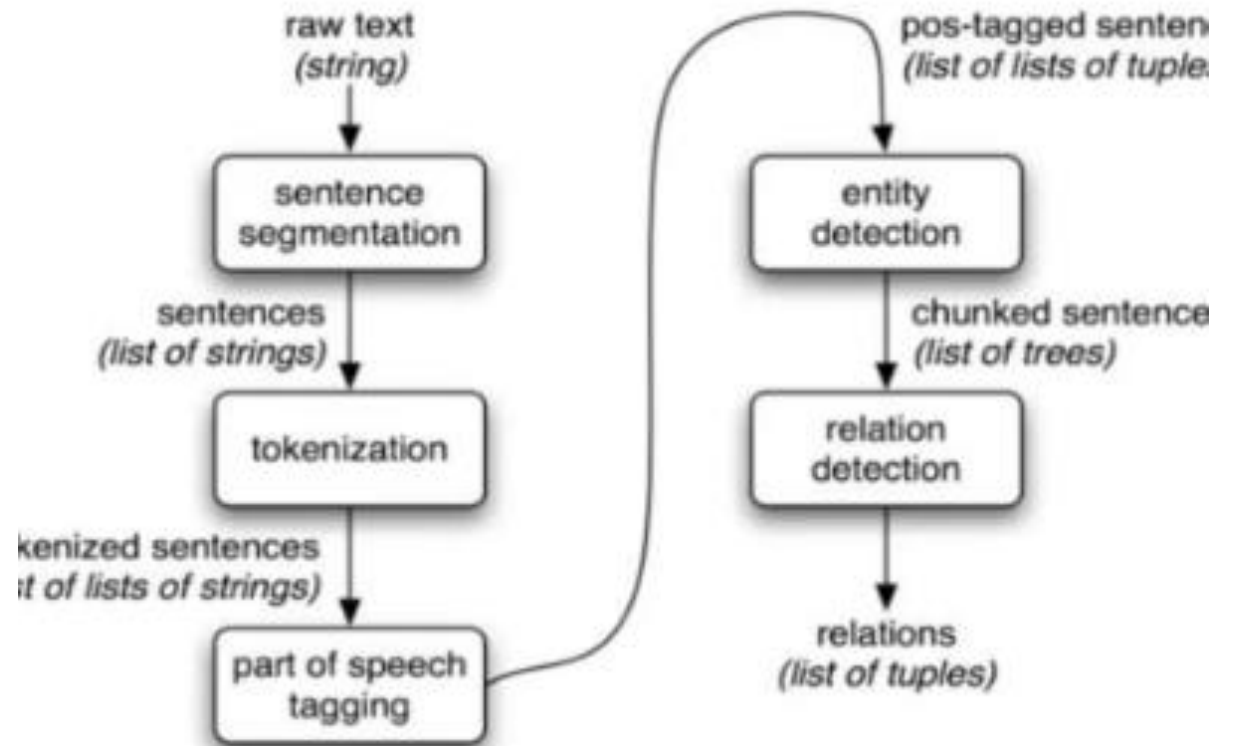


“ Natural Language Toolkit ”

- Leading platform for building Python program to work with human language data.
- NLP is a field of Computer Science, AI (ML), linguistics Processing
- Interaction between Computer and Human (natural) language
- IT provides suites of text processing libraries.

NLTK Libraries used in summarization

- Word Tokenizer
- Sentence Tokenizer
- Stopwords
- BeautifulSoup
- Numpy Library
- Tagging
- Parsing



Summarization

1

Segmentation

- Convert raw text into sentences
 - List of Strings
- Sentence Tokenizer

2

Tokenization

- Identifies the work token from given sentence
- Provides a list of token as output

3

Tagging

- Assigns appropriate part of speech tag to each word
- Extraction of nouns, adverbs, adjective.

4

Entity Detection

- Identification of predefined categories i.e., Person, location, quantities, organization.

5

Relation Detection

- Identifies the possible relation between two or more sentences
- Replacement with proper nouns

Conclusion & Scope

Automatic Text Summarization has been shown to be useful for natural language processing tasks such as questioning answering or text classification and other related fields of computer science such as information retrieval and access time for information searching will be improved.

Thank you...