

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes. Some nodes are solid blue circles, while others are white circles with blue outlines. The nodes are connected by thin, light gray lines, creating a mesh-like structure that extends towards the center of the slide.

# TEXT SUMMARIZATION USING NLP

Submitted By

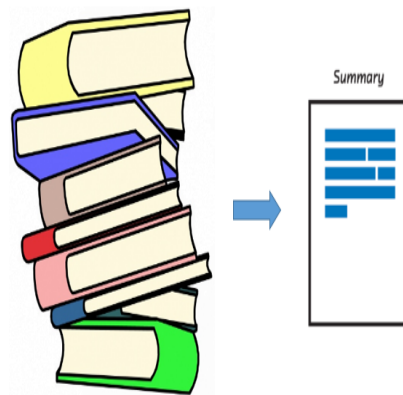
MD NADEEM SARWAR -20SCSE1010767  
HARSHIT MEHLAWAT -20SCSR1010943

---



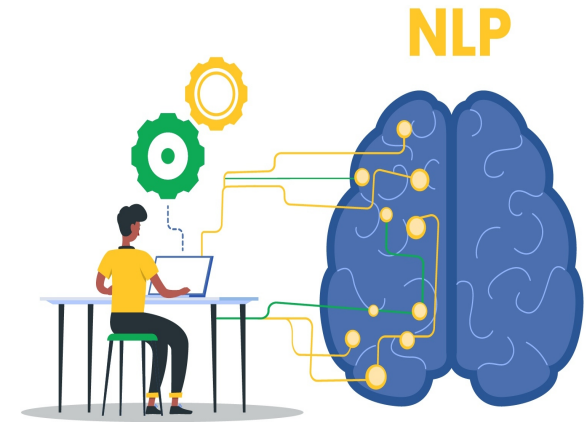
# A GENERAL INTRODUCTION TO TEXT SUMMARIZATION

- Text summarization is the problem of creating a short, accurate, and fluent summary of a longer text document.
- Automatic text summarization methods are greatly needed to address the ever-growing amount of text data available online to both better help discover relevant information and to consume relevant information faster.
- There is an enormous amount of textual material, and it is only growing every single day.
- Think of the internet, comprised of web pages, news articles, status updates, blogs and so much more. The data is unstructured and the best that we can do to navigate it is to use search and skim the results.



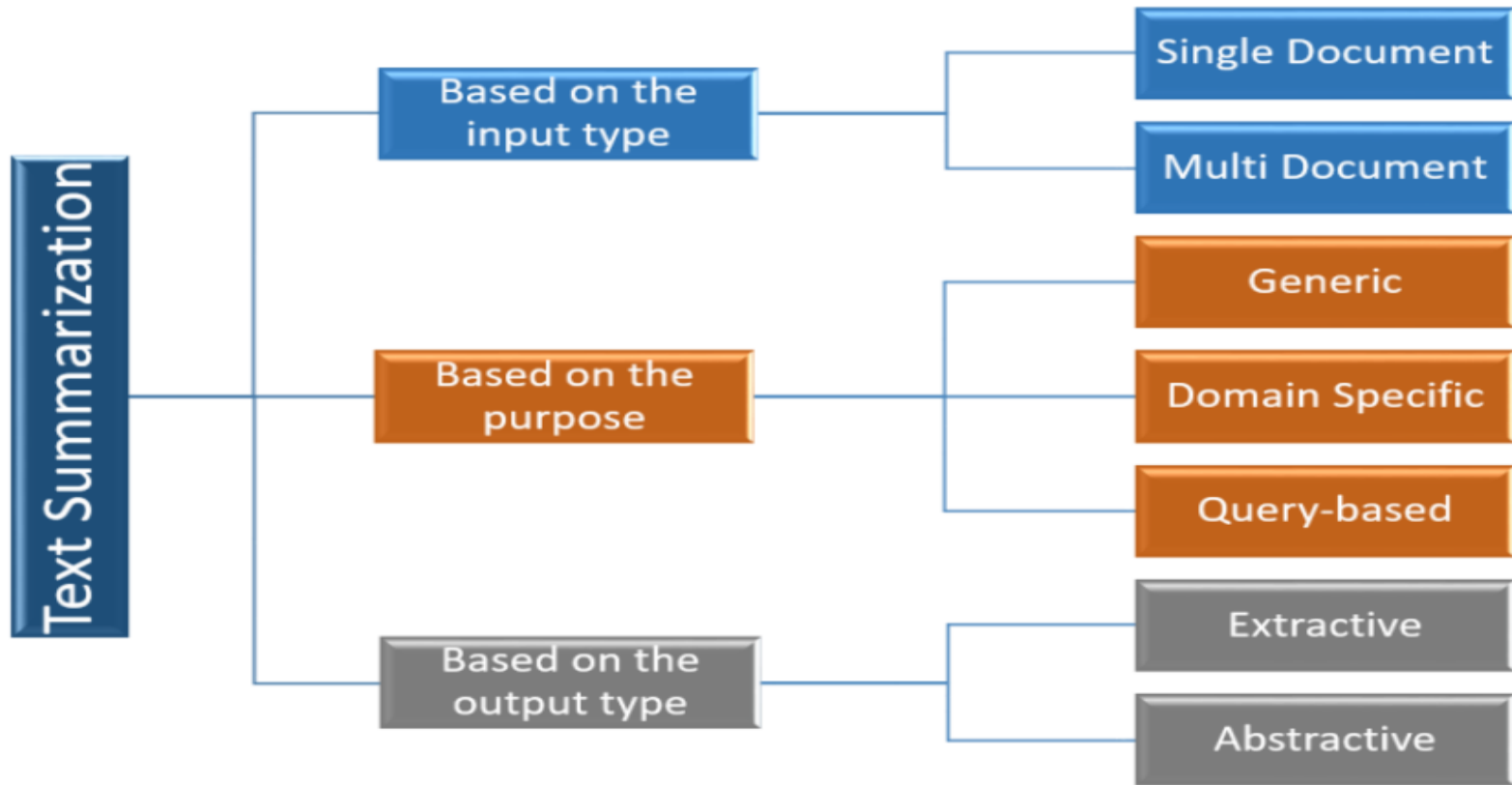
# WHAT IS NLP?

- Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken and written -- referred to as natural language.
- It is a component of artificial intelligence (AI)
- NLP combines computational linguistics—rule-based modeling of human language—with statistical, machine learning, and deep learning models.



# NEED OF TEXT SUMMARIZATION

- There is an enormous amount of textual material, and it is only growing every single day. Think of the internet, comprised of web pages, news articles, status updates, blogs and so much more.
- There is a great need to reduce much of this text data to shorter, focused summaries that capture the salient details, both so we can navigate it more effectively as well as check whether the larger documents contain the information that we are looking for.
- Text summarization reduces reading time, accelerates the process of researching for information, and increases the amount of information that can fit in an area.



# WHAT ARE THE 2 APPROACHES TO AUTOMATIC SUMMARIZATION

## TYPES

```
graph LR; A[TYPES] --> B[EXTRACTIVE SUMMARIZATION]; A --> C[ABSTRACTIVE SUMMARIZATION]
```

The diagram illustrates the two types of automatic summarization. A central box labeled 'TYPES' has two arrows pointing to two separate boxes. The top box is titled 'EXTRACTIVE SUMMARIZATION' and describes a method of selecting a subset of existing words, phrases, or sentences from the original text to form a summary. The bottom box is titled 'ABSTRACTIVE SUMMARIZATION' and describes a method that generates a sentence from a semantic representation and then uses natural language generation techniques to create a summary.

### EXTRACTIVE SUMMARIZATION

Extractive text summarization works by selecting a subset of existing words, phrases or sentences from the original text to form summary.

### ABSTRACTIVE SUMMARIZATION

Abstractive text summarization method generates a sentence from a semantic representation and then use natural language generation techniques to create a summary.


## A network diagram with nodes and connecting lines, overlaid with the CSS logo. The nodes are represented by circles of varying sizes and styles (some solid, some dashed). The connecting lines are thin and gray. The CSS logo is prominently displayed in the lower-left corner, featuring the letters 'CSS' in a bold, black, sans-serif font above a stylized blue and white graphic element.



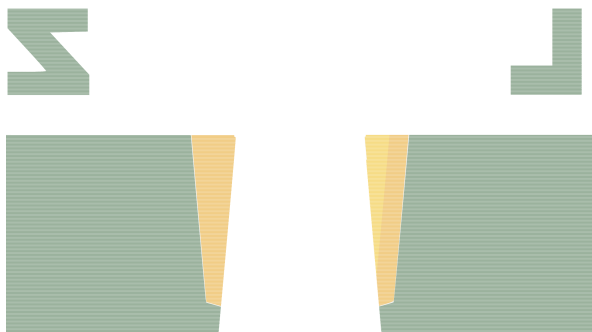
Ability of a computer program to understand human language as it is spoken and written



Hyper text markup language is the most basic building block of web. It defines the meaning & structure of web content.



Cascading style sheet is a stylesheet used to describe the presentation of a document written in XML or HTML



JavaScript is a scripting language that enables you to create dynamically updating content, control multimedia, animate images, and pretty much everything else.

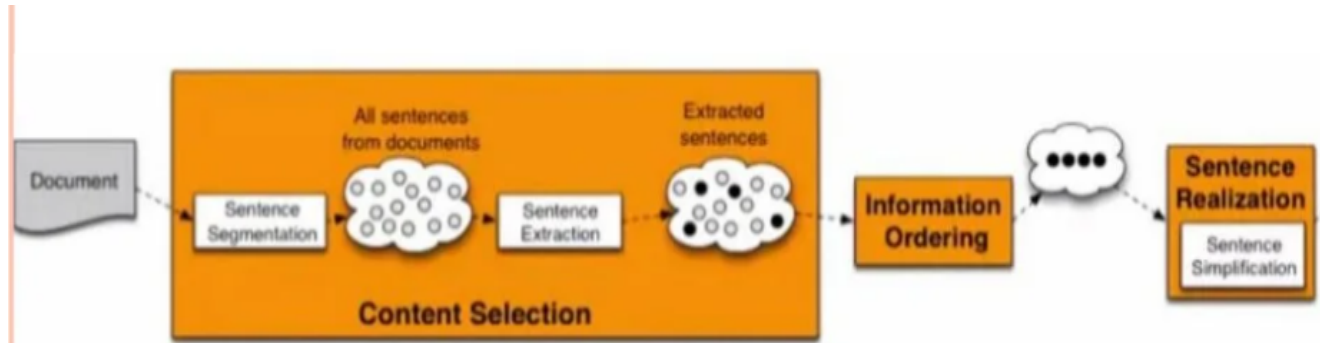


The summarization API allows you to summarize the meaning of a document extracting its most relevant sentences.



# SUMMARIZATION: 3 STAGES

- Content selection: choose sentences to extract from the documents.
- Information ordering: choose an order to place them in summary
- Sentence realization: clean up the sentence.



# ADVANTAGES OF TEXT SUMMARIZATION

- Summarizing reduces perusing time.
- While investigating reports, outlines make the determination process simpler.
- Summarization improves the adequacy of ordering.
- Text summarization are less biased than human summarizers.
- Personalized summaries are useful in question-answering systems.

**THANK YOU**

