Text Summarization using NLP

-Suhit Garg

-Pratyush Garg

-Shivam Goel

-Yash Sharma

Summarization has become a very helpful way of tackling the issue of data overburden in the 21st century. In this story, I will show you how you can create your personal text summarizer using Natural Language Processing (NLP) in Python.

**What is text summarization?**

It’s basically a task to generate an accurate summary while maintaining key information and not losing overall meaning.

There are two general types of summarization:

* **Abstractive summary** - generate new sentences from original text.
* **Extractive summary** - recognize important sentences and create a summary using those sentences.

**What will we use?**

We’ll use extractive summary because we can apply this method to many documents without having to do a lot of machine learning model training tasks.

Besides that, extractive summarization gives better summary outcome than abstractive summary, because abstractive summarization has to generate new sentences from the original text, which is a more difficult method than a data-driven approach to extract important sentences.

**Wok-flow**

Below is the workflow that we will be following

*import text*

*clean text and split into sentences*

*remove stop words*

*rank sentences*

*select top N sentences for summary*