**Text Summarization**

**Steps for Text Summarization:**

* Text cleaning
* Sentence tokenization
* Word tokenization
* Word-frequency table
* Summarization

**Detailed Approach:**

1. Created function **read\_file** to read the file data line by line, then splitting the text data where getting fullstop(.), and checking for all the alphabets present in the data.
2. Created function **sentence\_similarity** to check for similar sentences present, checked for unique characters in the sentences. Created vectors for all the words, calculated the cosine distance for the vectors.
3. Created function **gen\_sim\_matrix** to find out the similar words. Created similarity matrix by checking for the sentence similarity
4. Generated the summary by creating **generate\_summary** function, where created word frequency table, sentence similarity graph and summarized the text.

**Abstractive Text Summary:**

Abstractive text summarizers do not select sentences from the originally given text passage to create the summary. They produce a paraphrasing of the main contents of the given text, using a vocabulary set different from the original document. This is very similar to how we humans do, to summarize. We create a semantic representation of the document in our brains. We then pick words from our general vocabulary (the words we commonly use) that fit in the semantics, to create a short summary that represents all the points of the actual document.