```
import bs4
import urllib.request


data = urllib.request.urlopen('https://en.wikipedia.org/wiki/Classical_music')
data_read = data.read()


data_read = bs4.BeautifulSoup(data_read,'html.parser')


paras = data_read.find_all('p')


content = ''
for p in paras:
  content += p.text


content
```

```
'\nClassical music generally refers to the art music of the Western world, considered to be distinct from Western folk music or po
pular music traditions. It is sometimes distinguished as Western classical music, as the term "classical music" also applies to no
n-Western art music. Classical music is often characterized by formality and complexity in its musical form and harmonic organizat
ion,[1] particularly with the use of polyphony.[2] Since at least the ninth century it has been primarily a written tradition,[2]
spawning a sophisticated notational system, as well as accompanying literature in analytical, critical, historiographical, musicol
ogical and philosophical practices. A foundational component of Western culture, classical music is frequently seen from the persp
ective of individual or groups of composers, whose compositions, personalities and beliefs have fundamentally shaped its histor
```

```
from nltk.corpus.reader.tagged import word_tokenize
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
def dict_table(text_) -> dict :
      stop_words = set(stopwords.words("english"))
      words = word_tokenize(text_)
      stem = PorterStemmer()

      freq_table = dict()
      for wd in words:
        wd=stem.stem(wd)
        if wd in stop_words:
          continue
        if wd in freq_table:
          freq_table[wd] +=1
        else:
          freq_table[wd]=1

      return freq_table



import nltk
nltk.download('punkt')
nltk.download('stopwords')
from nltk.tokenize import word_tokenize , sent_tokenize
sentences  = sent_tokenize(content)
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
```

```
def calculate_sent_scores(sentences,freq_table) -> dict:
   sentence_weight = dict()

   for sentence in sentences:
        sentence_wordcount = (len(word_tokenize(sentence)))
        sentence_wordcount_without_stop_words = 0
        for word_weight in freq_table:
            if word_weight in sentence.lower():
                sentence_wordcount_without_stop_words += 1
                if sentence[:7] in sentence_weight:
                    sentence_weight[sentence[:7]] += freq_table[word_weight]
                else:
                    sentence_weight[sentence[:7]] = freq_table[word_weight]

        sentence_weight[sentence[:7]] = sentence_weight[sentence[:7]] /sentence_wordcount_without_stop_words

   return sentence_weight


def _calculate_average_score(sentence_weight) -> int:
```

```
    sum_values = 0
    for entry in sentence_weight:
        sum_values += sentence_weight[entry]


    average_score = (sum_values / len(sentence_weight))

    return average_score


def _get_article_summary(sentences, sentence_weight, threshold):
    sentence_counter = 0
    article_summary = ''

    for sentence in sentences:
        if sentence[:7] in sentence_weight and sentence_weight[sentence[:7]] >= (threshold):
            article_summary += " " + sentence
            sentence_counter += 1

    return article_summary


def _run_article_summary(article):


    frequency_table = dict_table(article)

    sentences = sent_tokenize(article)


    sentence_scores = calculate_sent_scores(sentences, frequency_table)


    threshold = _calculate_average_score(sentence_scores)


    article_summary = _get_article_summary(sentences, sentence_scores, 1.5 * threshold)

    return article_summary



if __name__ == '__main__':
    summary_results = _run_article_summary(content)
    print(summary_results)
```

```
ırdization of descriptions and specifications of instruments, as well as instruction in their use. He also composed Euridice, the fir
```

Colab paid products - Cancel contracts here

✓ 0s completed at 11:26 PM