```
!pip install rouge
import pandas as pd
import numpy as np
from rouge import Rouge

def rouge_scores(hypothesis, reference):
    rouge = Rouge()
    scores = rouge.get_scores(hypothesis, reference)
    return scores

import nltk
nltk.download('stopwords')
nltk.download('punkt')
from nltk.tokenize import word_tokenize, sent_tokenize
from nltk.corpus import stopwords
def solve(text):
  stopwords1 = set(stopwords.words("english"))
  words = word_tokenize(text)
  freqTable = {}
  for word in words:
    word = word.lower()
    if word in stopwords1:
      continue
    if word in freqTable:
      freqTable[word] += 1
    else:
      freqTable[word] = 1

  sentences = sent_tokenize(text)
  sentenceValue = {}
  for sentence in sentences:
    for word, freq in freqTable.items():
      if word in sentence.lower():
        if sentence in sentenceValue:
          sentenceValue[sentence] += freq
        else:
          sentenceValue[sentence] = freq
  sumValues = 0
  for sentence in sentenceValue:
    sumValues += sentenceValue[sentence]
  average = int(sumValues / len(sentenceValue))

  summary = ''
  for sentence in sentences:
    if (sentence in sentenceValue) and(sentenceValue[sentence] > (1.2 * average)):
      summary += "" + sentence
  return summary
data = '''As per the survey, stemming in used as a pre-processing step in most of the existing research, although the generated stem may not I
also identified from the survey that in most cases, the Natural Language Toolkit (NLTK) alone was used for conducting pre-processing tasks, a
lemmatization and Part-of-Speech (POS) tagging accurately. In the proposed model, Stanford CoreNLP [11] is used alongside NLTK as it yields b
is a stage that is added in the proposed approach to mitigate
overfitting.A combination of various features, that have been considered in separate experiments in existing research and have proven to be e
of Support Vector Machine (SVM), K-Nearest Neighbour (KNN) and Decision Tree algorithms is compared and the generated summary is also convert
s=solve(data)
print("\n\n\n",s)
print("\n\nRouge scores:",rouge_scores(data,s))
```

```
Requirement already satisfied: rouge in /usr/local/lib/python3.10/dist-packages (1.0.1)
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (from rouge) (1.16.0)


    It was
    also identified from the survey that in most cases, the Natural Language Toolkit (NLTK) alone was used for conducting pre-processing tas
    lemmatization and Part-of-Speech (POS) tagging accurately.


    Rouge scores: [{'rouge-1': {'r': 1.0, 'p': 0.24444444444444444, 'f': 0.39285713970025515}, 'rouge-2': {'r': 1.0, 'p': 0.17346938775510204
    [nltk_data] Downloading package stopwords to /root/nltk_data...
    [nltk_data]   Package stopwords is already up-to-date!
    [nltk_data] Downloading package punkt to /root/nltk_data...
    [nltk_data]   Package punkt is already up-to-date!
```