!pip install nltk -U

!pip install bs4 -U

import nltk

nltk.download('stopwords')

nltk.download('punkt')

nltk.download('wordnet')

nltk.download('averaged\_perceptron\_tagger')

import nltk

para='Engineering is a field that encompasses the application of scientific and mathematical principles to design, develop, and improve structures, machines, systems, and processes. It plays a crucial role in shaping the world we live in, from the buildings we inhabit to the technology we use daily. Engineers are problem solvers, using their creativity and technical expertise to address challenges and create innovative solutions. Whether its designing sustainable infrastructure, developing new medical devices, or optimizing industrial processes, engineering is at the forefront of driving progress and improving quality of life.'

print(para)

para.split()

from nltk.tokenize import sent\_tokenize

sent=sent\_tokenize(para)

sent[1]

from nltk.tokenize import word\_tokenize

words=word\_tokenize(para)

words

from nltk.corpus import stopwords

swords=stopwords.words('english')

swords

x=[word for word in words if word not in swords]

x

x=[word for word in words if word.lower() not in swords]

x

from nltk.stem import PorterStemmer

ps=PorterStemmer()

ps.stem('hashing')

y=[ps.stem(word)for word in x]

y

from nltk.stem import WordNetLemmatizer

wnl=WordNetLemmatizer()

nltk.download('omw-1.4')

wnl.lemmatize('hashing',pos='v')

print(ps.stem('went'))

print(wnl.lemmatize('gone',pos='v'))

z=[wnl.lemmatize(word,pos='v') for word in x]

z

import string

string.punctuation

t=[word for word in words if word not in string.punctuation]

t

from nltk import pos\_tag

pos\_tag(t)