import nltk

nltk.download('punkt')

nltk.download('wordnet')

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

nltk.download('stopwords')

from nltk import sent\_tokenize

from nltk import word\_tokenize

from nltk.corpus import stopwords

text='Real madrid is set to win the UCL for the season . Benzema might win Balon dor . Salah might be the runner up'

tokens\_sents = nltk.sent\_tokenize(text)

print(tokens\_sents)

tokens\_words = nltk.word\_tokenize(text)

print(tokens\_words)

from nltk.stem import PorterStemmer

from nltk.stem.snowball import SnowballStemmer

from nltk.stem import LancasterStemmer

stem=[]

for i in tokens\_words:

ps = PorterStemmer()

stem\_word= ps.stem(i)

stem.append(stem\_word)

print(stem)

#Lemmatization

import nltk

from nltk.stem import WordNetLemmatizer

lemmatizer = WordNetLemmatizer()

lemmatized\_output = ' '.join([lemmatizer.lemmatize(w) for w in stem])

print(lemmatized\_output)

leme=[]

for i in stem:

lemetized\_word=lemmatizer.lemmatize(i)

leme.append(lemetized\_word)

print(leme)

#Part of SPeech Tagging

print("Parts of Speech: ",nltk.pos\_tag(leme))

#Stop Word

sw\_nltk = stopwords.words('english')

print(sw\_nltk)

words = [word for word in text.split() if word.lower() not in sw\_nltk]

new\_text = " ".join(words)

print(new\_text)