def tokenize\_text(text):

# Split the text by spaces

tokens = []

for word in text.split():

# Remove leading and trailing punctuation

clean\_word = word.strip(",.?!:;'\"()[]{}")

if clean\_word: # Avoid adding empty strings

tokens.append(clean\_word)

return tokens

sample\_inputs = [ "Zoro is a swordman."

, "He loves war.",

"His favourite sword style is three sword style."]

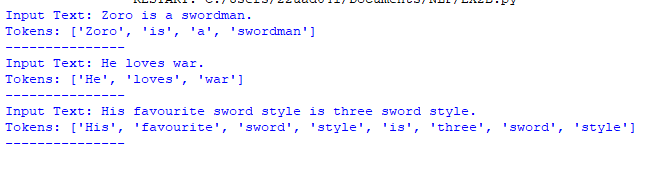
for input\_text in sample\_inputs:

print("Input Text:", input\_text)

tokens = tokenize\_text(input\_text)

print("Tokens:", tokens)

print("---------------")



import nltk

def tokenize\_text(text):

tokens = nltk.word\_tokenize(text)

return tokens

def split\_sentences(text):

sentences = nltk.sent\_tokenize(text)

return sentences

sample\_inputs = ["Zoro is a swordman.

He loves war.

His favourite sword style is three sword style."]

for input\_text in sample\_inputs:

print("Input Text:", input\_text)

sentences = split\_sentences(input\_text)

print("Sentences:", sentences)

for sentence in sentences:

tokens = tokenize\_text(sentence)

print("Tokens for sentence:", sentence)

print("Tokens:", tokens)

print("---------------")

