###  Assignment No 4 ###

"""Assignment Title : Implement Bi-gram, Tri-gram word sequence and its count in text input

data using NLTK library"""

from nltk import ngrams

from nltk.util import ngrams

#unigram model

n = 1

sentence = 'Earth is the third planet from the Sun in our solar system and the only known celestial body to support life. With a diverse range of ecosystems, it is home to a vast array of plant and animal species, including humans.'

unigrams = ngrams(sentence.split(), n)

print(f"\n\*\*\*\*\*\*\*\*\*\*\*   UNIGRAM    \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

for item in unigrams:

    print(item)

#bigram model

n = 2

sentence = 'Earth is the third planet from the Sun in our solar system and the only known celestial body to support life. With a diverse range of ecosystems, it is home to a vast array of plant and animal species, including humans.'

unigrams = ngrams(sentence.split(), n)

print(f"\n\*\*\*\*\*\*\*\*\*\*\*   BIGRAM    \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

for item in unigrams:

    print(item)

#trigram model

n = 3

sentence = 'Earth is the third planet from the Sun in our solar system and the only known celestial body to support life. With a diverse range of ecosystems, it is home to a vast array of plant and animal species, including humans.'

unigrams = ngrams(sentence.split(), n)

print(f"\n\*\*\*\*\*\*\*\*\*\*\*   TRIGRAM    \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

for item in unigrams:

    print(item)

'''

\*\*\*\*\*\*\*\*\*\*\*\*    OUTPUT    \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\*\*\*\*\*\*\*\*   UNIGRAM    \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

('Earth',)

('is',)

('the',)

('third',)

('planet',)

('from',)

('the',)

('Sun',)

('in',)

('our',)

('solar',)

('system',)

('and',)

('the',)

('only',)

('known',)

('celestial',)

('body',)

('to',)

('support',)

('life.',)

('With',)

('a',)

('diverse',)

('range',)

('of',)

('ecosystems,',)

('it',)

('is',)

('home',)

('to',)

('a',)

('vast',)

('array',)

('of',)

('plant',)

('and',)

('animal',)

('species,',)

('including',)

('humans.',)

\*\*\*\*\*\*\*\*\*\*\*   BIGRAM    \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

('Earth', 'is')

('is', 'the')

('the', 'third')

('third', 'planet')

('planet', 'from')

('from', 'the')

('the', 'Sun')

('Sun', 'in')

('in', 'our')

('our', 'solar')

('solar', 'system')

('system', 'and')

('and', 'the')

('the', 'only')

('only', 'known')

('known', 'celestial')

('celestial', 'body')

('body', 'to')

('to', 'support')

('support', 'life.')

('life.', 'With')

('With', 'a')

('a', 'diverse')

('diverse', 'range')

('range', 'of')

('of', 'ecosystems,')

('ecosystems,', 'it')

('it', 'is')

('is', 'home')

('home', 'to')

('to', 'a')

('a', 'vast')

('vast', 'array')

('array', 'of')

('of', 'plant')

('plant', 'and')

('and', 'animal')

('animal', 'species,')

('species,', 'including')

('including', 'humans.')

\*\*\*\*\*\*\*\*\*\*\*   TRIGRAM    \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

('Earth', 'is', 'the')

('is', 'the', 'third')

('the', 'third', 'planet')

('third', 'planet', 'from')

('planet', 'from', 'the')

('from', 'the', 'Sun')

('the', 'Sun', 'in')

('Sun', 'in', 'our')

('in', 'our', 'solar')

('our', 'solar', 'system')

('solar', 'system', 'and')

('system', 'and', 'the')

('and', 'the', 'only')

('the', 'only', 'known')

('only', 'known', 'celestial')

('known', 'celestial', 'body')

('celestial', 'body', 'to')

('body', 'to', 'support')

('to', 'support', 'life.')

('support', 'life.', 'With')

('life.', 'With', 'a')

('With', 'a', 'diverse')

('a', 'diverse', 'range')

('diverse', 'range', 'of')

('range', 'of', 'ecosystems,')

('of', 'ecosystems,', 'it')

('ecosystems,', 'it', 'is')

('it', 'is', 'home')

('is', 'home', 'to')

('home', 'to', 'a')

('to', 'a', 'vast')

('a', 'vast', 'array')

('vast', 'array', 'of')

('array', 'of', 'plant')

('of', 'plant', 'and')

('plant', 'and', 'animal')

('and', 'animal', 'species,')

('animal', 'species,', 'including')

('species,', 'including', 'humans.')

'''