Name: Manya Madhu

Roll No:17

Batch: MCA-B

Date: 14-11-2022

### **DATA SCIENCE LAB**

### **Experiment No.: 12**

# <u>Ai</u>m

N grams

## **Procedure**

a)

```
import nltk
from nltk.util import ngrams
text = "this is a very good book to study";
Ngrams = ngrams(sequence=nltk.wordpunct_tokenize(text), n=3)
for grams in Ngrams:
    print(grams)
```

### Output:

```
('this', 'is', 'a')
('is', 'a', 'very')
('a', 'very', 'good')
('very', 'good', 'book')
('good', 'book', 'to')
('book', 'to', 'study')
```

### b)

```
import requests
from bs4 import BeautifulSoup
import csv

URL = "http://www.values.com/inspirational-quotes"
r = requests.get(URL)

soup = BeautifulSoup(r.content, 'html5lib')

print(soup.prettify())
```

#### Output

```
<!DOCTYPE html>
<html class="no-js" dir="ltr" lang="en-US">
 <meta charset="utf-8"/>
 <meta content="text/html; charset=utf-8" http-equiv="content-type"/>
 <meta content="IE=edge" http-equiv="X-UA-Compatible"/>
 <meta content="width=device-width,initial-scale=1.0" name="viewport"/>
  Inspirational Quotes - Motivational Quotes - | The Foundation for a
Better Life
 </title>
 <meta content="Find the perfect quotation from our hand-picked collection</pre>
of inspiring quotes by hundreds of authors." name="description"/>
  <meta content="pass, it, on, passiton, values, kindness"</pre>
name="keywords"/>
 <meta content="The Foundation for a Better Life"</pre>
name="twitter:site name"/>
 <meta content="@passiton values" name="twitter:site"/>
 <meta content="summary" name="twitter:card"/>
 <meta content="Thank you for visiting." name="twitter:description"/>
 <meta content="https://www.passiton.com/passiton fbl.jpg"</pre>
name="twitter:image"/>
 <meta content="https://www.passiton.com/inspirational-quotes"</pre>
property="og:url"/>
 <meta content="The Foundation for a Better Life"</pre>
property="og:site name"/>
 <meta content="Welcome to PassItOn.com" property="og:title"/>
 <meta content="https://www.passiton.com/passiton fbl.jpg"</pre>
property="og:image"/>
 <meta content="Thank you for visiting." property="og:description"/>
 <meta content="website" property="og:type"/>
 <link href="/apple-touch-icon.png" rel="apple-touch-icon"</pre>
sizes="180x180"/>
 <link href="/favicon-32x32.png" rel="icon" sizes="32x32"</pre>
type="image/png"/>
 <link href="/favicon-16x16.png" rel="icon" sizes="16x16"</pre>
type="image/png"/>
 <link href="/site.webmanifest" rel="manifest"/>
 <link color="#c8102e" href="/safari-pinned-tab.svg" rel="mask-icon"/>
 <meta content="#ffffff" name="msapplication-TileColor"/>
 <meta content="#ffffff" name="theme-color"/>
 <link crossorigin="anonymous"</pre>
href="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.
css" integrity="sha384-
qqOyR0iXCbMQv3Xipma34MD+dH/1fQ784/j6cY/iJTQUOhcWr7x9JvoRxT2MZw1T"
rel="stylesheet"/>
 <link data-turbo-track="reload" href="/assets/application-</pre>
e8e20e255c69ae39d804fef20f59772f15fb1bc96bf62568750ac089c781848a.css"
rel="stylesheet"/>
 <meta content="authenticity token" name="csrf-param"/>
  <meta
content="9VHqfcfQZnOvp SCSW1FIX8doR3IKA 4f4jfYi q2uSjnhPwRCA6AFmwhvCGJM1VfA
x--FgRVlcCBlhOGjvJGQ" name="csrf-token"/>
 <!-- Global site tag (gtag.js) - Google Analytics -->
 <script async="" src="https://www.googletagmanager.com/gtag/js?id=UA-</pre>
1179606-29">
 </script>
  <script>
```