DATA SCIENCE LAB

Name: SWETHA PRAKASH

Roll No: 46

Batch: B

Date: 14/11/2022

Experiment No.: 17

Aim

Implement ngrams and BeautifulSoup algorithm using python.

Procedure

ngrams

```
import nltk
from nltk.util import ngrams
text = "I have no idea what you are talking about";
Ngrams = ngrams(sequence=nltk.wordpunct_tokenize(text), n=3)
for grams in Ngrams:
print(grams)
```

```
('I', 'have', 'no')
('have', 'no', 'idea')
('no', 'idea', 'what')
('idea', 'what', 'you')
('what', 'you', 'are')
('you', 'are', 'talking')
('are', 'talking', 'about')
```

BeautifulSoup

```
import requests
from bs4 import BeautifulSoup
import csv

URL = "https://www.aes.ajce.in"
r = requests.get(URL)
soup = BeautifulSoup(r.content, 'html5lib')
print(soup.prettify())
```

```
<html lang="en">
 <head>
  <meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
  <meta content="width=device-width, initial-scale=1.0" name="viewport"/>
  <title>
   Academic Enterprise Solutions (AES)- Staff / Students / Parents Login for Amal Jyothi College of Engineering
  </title>
  <link href="aesicon.ico" rel="icon" type="image/ico"/>
  <!--STYLESHEET-->
  <!--=================================================-->
  <!--Open Sans Font [ OPTIONAL ]-->
  <link href="https://fonts.googleapis.com/css?family=Open+Sans:400,300,600,700" rel="stylesheet" type="text/css"/>
  <!--Bootstrap Stylesheet [ REQUIRED ]-->
  <link href="css/bootstrap.min.css" rel="stylesheet"/>
  <!--Nifty Stylesheet [ REQUIRED ]-->
  <link href="css/nifty.min.css" rel="stylesheet"/>
  <!--Nifty Premium Icon [ DEMONSTRATION ]-->
  <link href="css/demo/nifty-demo-icons.min.css" rel="stylesheet"/>
  <!--Demo [ DEMONSTRATION ]-->
  <link href="css/demo/nifty-demo.min.css" rel="stylesheet"/>
  <!--Magic Checkbox [ OPTIONAL ]-->
  <link href="plugins/magic-check/css/magic-check.min.css" rel="stylesheet"/>
  <link href="plugins/font-awesome/css/font-awesome.min.css" rel="stylesheet"/>
  <link href="/plugins/animate-css/animate.min.css" rel="stylesheet"/>
  <!--JAVASCRIPT-->
```