

181b069

INFORMATION RETRIEVAL LAB

BALAJI G

CODE IN PYTHON 3.9:

```
Lab 1.py - C:\Users\Admin\Desktop\Lab 1.py (3.9.1)
File Edit Format Run Options Window Help
## 181B069 BALAJI G
from docx import Document
import pandas as pd
document = Document()
document = Document(r'C:\Users\Admin\Downloads\Assignment task 1.docx')
table=document.tables[0]
data = [[cell.text for cell in row.cells] for row in table.rows]
df = pd.DataFrame(data)
print(df)
query=input()
a=[]
for col in df:
    a.append(df[col].str.contains(query))
for ind in df.index:
    if a[1][ind] == True:
        print(df[0][ind])
```

QUERY 1: Natural language processing

OUTPUT :

```
Lab 1.py - C:\Users\Admin\Desktop\Lab 1.py (3.9.1)
File Edit Format Run Options Window Help
## 181B069 BALAJI G
from docx import Document
import pandas as pd
document = Document()
document = Document(r'C:\Users\Admin\Downloads\Assignment task 1.docx')
table=document.tables[0]
data = [[cell.text for cell in row.cells] for row in table.rows]
df = pd.DataFrame(data)
print(df)
query=input()
a=[]
for col in df:
    a.append(df[col].str.contains(query))
for ind in df.index:
    if a[1][ind] == True:
        print(df[0][ind])

Python 3.9.1 (tags/v3.9.1:1e5d33e, Dec 7 2020, 17:08:21) [MSC v.1927 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
===== RESTART: C:\Users\Admin\Desktop\Lab 1.py =====
0          1
0  Document Number      Document Text
1  Statistical approaches to processing natural l...
2  This book offers a highly accessible introduct...
3  Class-tested and coherent, this groundbreaking...
4  Class-tested and coherent, this textbook teach...
5  An information retrieval (IR) system is design...
6  Class-tested and coherent, this textbook teach...
7  Processing multimedia content has emerged as a...
8  Through a series of recent breakthroughs, deep...
9  Learn how to use the latest techniques, best p...
10 Find out how HTML lays the foundation of web d...
natural language processing
1
2
>>> |
```

QUERY 2: Information Retrieval

OUTPUT:

```
IDLE Shell 3.9.1
File Edit Shell Debug Options Window Help
6      6 Class-tested and coherent, this textbook teach...
7      7 Processing multimedia content has emerged as a...
8      8 Through a series of recent breakthroughs, deep...
9      9 Learn how to use the latest techniques, best p...
10     10 Find out how HTML lays the foundation of web d...
natural language processing
1
2
>>>
===== RESTART: C:\Users\Admin\Desktop\Lab 1.py =====
0      0      1
0 Document Number      Document Text
1      1 Statistical approaches to processing natural l...
2      2 This book offers a highly accessible introduct...
3      3 Class-tested and coherent, this groundbreaking...
4      4 Class-tested and coherent, this textbook teach...
5      5 An information retrieval (IR) system is design...
6      6 Class-tested and coherent, this textbook teach...
7      7 Processing multimedia content has emerged as a...
8      8 Through a series of recent breakthroughs, deep...
9      9 Learn how to use the latest techniques, best p...
10     10 Find out how HTML lays the foundation of web d...
information retrieval
1
3
4
5
6
>>>
```

QUERY 3: Book text

OUTPUT :

```
IDLE Shell 3.9.1
File Edit Shell Debug Options Window Help
4      4 Class-tested and coherent, this textbook teach...
5      5 An information retrieval (IR) system is design...
6      6 Class-tested and coherent, this textbook teach...
7      7 Processing multimedia content has emerged as a...
8      8 Through a series of recent breakthroughs, deep...
9      9 Learn how to use the latest techniques, best p...
10     10 Find out how HTML lays the foundation of web d...
information retrieval
1
3
4
5
6
>>>
===== RESTART: C:\Users\Admin\Desktop\Lab 1.py =====
0      0      1
0 Document Number      Document Text
1      1 Statistical approaches to processing natural l...
2      2 This book offers a highly accessible introduct...
3      3 Class-tested and coherent, this groundbreaking...
4      4 Class-tested and coherent, this textbook teach...
5      5 An information retrieval (IR) system is design...
6      6 Class-tested and coherent, this textbook teach...
7      7 Processing multimedia content has emerged as a...
8      8 Through a series of recent breakthroughs, deep...
9      9 Learn how to use the latest techniques, best p...
10     10 Find out how HTML lays the foundation of web d...
book text
>>>
```

QUERY 4 : Book

OUTPUT :

```
IDLE Shell 3.9.1
File Edit Shell Debug Options Window Help
7 Processing multimedia content has emerged as a...
8 Through a series of recent breakthroughs, deep...
9 Learn how to use the latest techniques, best p...
10 Find out how HTML lays the foundation of web d...
book text
>>>
===== RESTART: C:\Users\Admin\Desktop\Lab 1.py =====
0 Document Number Document Text
1 1 Statistical approaches to processing natural l...
2 2 This book offers a highly accessible introduct...
3 3 Class-tested and coherent, this groundbreaking...
4 4 Class-tested and coherent, this textbook teach...
5 5 An information retrieval (IR) system is design...
6 6 Class-tested and coherent, this textbook teach...
7 7 Processing multimedia content has emerged as a...
8 8 Through a series of recent breakthroughs, deep...
9 9 Learn how to use the latest techniques, best p...
10 10 Find out how HTML lays the foundation of web d...
book
1
2
3
4
5
6
7
8
>>>
```

QUERY 5 : Web HTML

OUTPUT :

```
IDLE Shell 3.9.1
File Edit Shell Debug Options Window Help
0 Document Number Document Text
1 1 Statistical approaches to processing natural l...
2 2 This book offers a highly accessible introduct...
3 3 Class-tested and coherent, this groundbreaking...
4 4 Class-tested and coherent, this textbook teach...
5 5 An information retrieval (IR) system is design...
6 6 Class-tested and coherent, this textbook teach...
7 7 Processing multimedia content has emerged as a...
8 8 Through a series of recent breakthroughs, deep...
9 9 Learn how to use the latest techniques, best p...
10 10 Find out how HTML lays the foundation of web d...
web HTML
>>>
===== RESTART: C:\Users\Admin\Desktop\Lab 1.py =====
0 Document Number Document Text
1 1 Statistical approaches to processing natural l...
2 2 This book offers a highly accessible introduct...
3 3 Class-tested and coherent, this groundbreaking...
4 4 Class-tested and coherent, this textbook teach...
5 5 An information retrieval (IR) system is design...
6 6 Class-tested and coherent, this textbook teach...
7 7 Processing multimedia content has emerged as a...
8 8 Through a series of recent breakthroughs, deep...
9 9 Learn how to use the latest techniques, best p...
10 10 Find out how HTML lays the foundation of web d...
Web HTML
>>>
```

QUERY 6 : Text theory

OUTPUT :

```
IDLE Shell 3.9.1
File Edit Shell Debug Options Window Help
0
1
2
3
4
5
6
7
8
9
10
Web HTML
>>>
===== RESTART: C:\Users\Admin\Desktop\Lab 1.py =====
0
1
2
3
4
5
6
7
8
9
10
text theory
>>>
```

The image shows a screenshot of an IDLE Shell 3.9.1 window. The window displays a list of 10 documents with their numbers and titles. The titles are truncated. The list is as follows:

Document Number	Document Text
1	Statistical approaches to processing natural l...
2	This book offers a highly accessible introduct...
3	Class-tested and coherent, this groundbreaking...
4	Class-tested and coherent, this textbook teach...
5	An information retrieval (IR) system is design...
6	Class-tested and coherent, this textbook teach...
7	Processing multimedia content has emerged as a...
8	Through a series of recent breakthroughs, deep...
9	Learn how to use the latest techniques, best p...
10	Find out how HTML lays the foundation of web d...

The window also shows a prompt for 'Web HTML' and a 'RESTART' message. The user has entered 'text theory' and the prompt is '>>>'. The taskbar at the bottom shows several open applications, including 'Lab assignment 16...', '181b069_lab1 - Mic...', 'Lab 1.py - C:\Users\...', and 'IDLE Shell 3.9.1'. The system clock shows '12:27 AM 17-01-2021'.

```
from docx import Document
import pandas as pd
document = Document()
document = Document(r'C:\Users\Admin\Downloads\Assignment task 1.docx')
table=document.tables[o]
data = [[cell.text for cell in row.cells] for row in table.rows]
df = pd.DataFrame(data)
print(df)
query=input()
a=[]
for col in df:
    a.append(df[col].str.contains(query))
for ind in df.index:
    if a[1][ind] == True:
        print(df[o][ind])
```