



IR ASSIGNMENT

Personal Repository of Web Content

AADILA JASMIN
2019A7PS0801U

I. Data Preparation Step:

The keywords used in this project are related to my research paper on “Facial Mask Detection and Identification of Individual during the Covid-19 Pandemic”.

5 keywords were selected:

1) Computer Vision

2) Deep Learning

3) Face Mask Detection

4) Haar Cascade Algorithm

5) K nearest neighbor

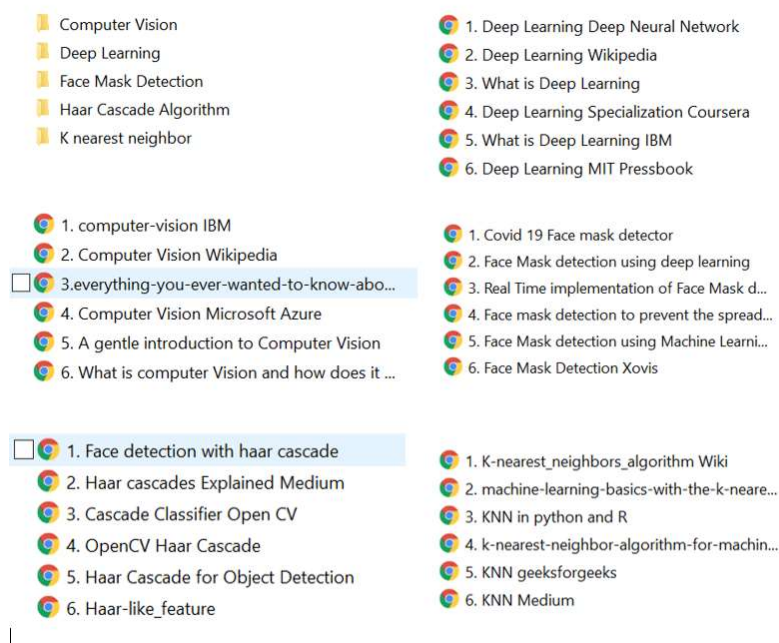


Figure 1- Five Different folders for keywords and the 6 files in each of the folders

After selection of the 5 keywords, the top 6 files were found using the Google Search Engine and saved on **MongoDB database manually with the keyword, file name and path.**

Documents Aggregations Schema Explain Plan Indexes Validation

FILTER { field: 'value' }

ADD DATA **VIEW**

```
_id: ObjectId("61685ca0018cce4cd877e7bc")
keyword: "Computer Vision"
files: Object
  1. computer-vision IBM: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Computer Vision\1. compute..."
  2. Computer Vision Wiki: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Computer Vision\2. Compute..."
  3. everything-you-ever-w: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Computer Vision\3. everythi..."
  4. Computer Vision Micr: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Computer Vision\4. Compute..."
  5. A gentle introductio: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Computer Vision\5. A gentl..."
  6. What is computer Vis: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Computer Vision\6. What is..."
```

```
_id: ObjectId("61685ccd018cce4cd877e7be")
keyword: "Deep Learning"
files: Object
  1. Deep Learning Deep N: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Deep Learning\1. Deep Lear..."
  2. Deep Learning Wikip: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Deep Learning\2. Deep Lear..."
  3. What is Deep Learning: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Deep Learning\3. What is D..."
  4. Deep Learning Specia: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Deep Learning\4. Deep Lear..."
  5. What is Deep Learnin: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Deep Learning\5. What is D..."
  6. Deep Learning MIT Pr: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Deep Learning\6. Deep Lear..."
```

```
_id: ObjectId("61685cee018cce4cd877e7c0")
keyword: "Face Mask Detection"
files: Object
  1. Covid 19 Face mask d: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\1. Cov..."
  2. Face Mask detection : "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\2. Fac..."
  3. Real Time implementa: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\3. Rea..."
  4. Face mask detection : "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\4. Fac..."
```

Figure 2- Storing the file with the keyword, filename and path in MongoDB

Documents Aggregations Schema Explain Plan Indexes Validation

FILTER { field: 'value' }

ADD DATA **VIEW**

```
files: Object
  1. Covid 19 Face mask d: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\1. Cov..."
  2. Face Mask detection : "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\2. Fac..."
  3. Real Time implementa: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\3. Rea..."
  4. Face mask detection : "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\4. Fac..."
  5. Face Mask detection : "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\5. Fac..."
  6. Face Mask Detection : "C:\Users\aadil\OneDrive\Desktop\lab2\files\Face Mask Detection\6. Fac..."
```

```
_id: ObjectId("61685d05018cce4cd877e7c2")
keyword: "Haar Cascade Algorithm"
files: Object
  1. Face detection with : "C:\Users\aadil\OneDrive\Desktop\lab2\files\Haar Cascade Algorithm\1. ..."
  2. Haar cascades Explai: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Haar Cascade Algorithm\2. ..."
  3. Cascade Classifier O: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Haar Cascade Algorithm\3. ..."
  4. OpenCV Haar Cascade: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Haar Cascade Algorithm\4. ..."
  5. Haar Cascade for Obj: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Haar Cascade Algorithm\5. ..."
  6. Haar-like_feature: "C:\Users\aadil\OneDrive\Desktop\lab2\files\Haar Cascade Algorithm\6. ..."
```

```
_id: ObjectId("61685d1f018cce4cd877e7c4")
keyword: "K nearest neighbor"
files: Object
  K-nearest_neighbors_al: "C:\Users\aadil\OneDrive\Desktop\lab2\files\K nearest neighbor\1. K-ne..."
  machine-learning-basics: "C:\Users\aadil\OneDrive\Desktop\lab2\files\K nearest neighbor\2. mach..."
  KNN in python and R: "C:\Users\aadil\OneDrive\Desktop\lab2\files\K nearest neighbor\3. KNN ..."
  k-nearest-neighbor-algo: "C:\Users\aadil\OneDrive\Desktop\lab2\files\K nearest neighbor\4. k-ne..."
  KNN geeksforgeeks: "C:\Users\aadil\OneDrive\Desktop\lab2\files\K nearest neighbor\5. KNN ..."
  KNN Medium: "C:\Users\aadil\OneDrive\Desktop\lab2\files\K nearest neighbor\6. KNN ..."
```

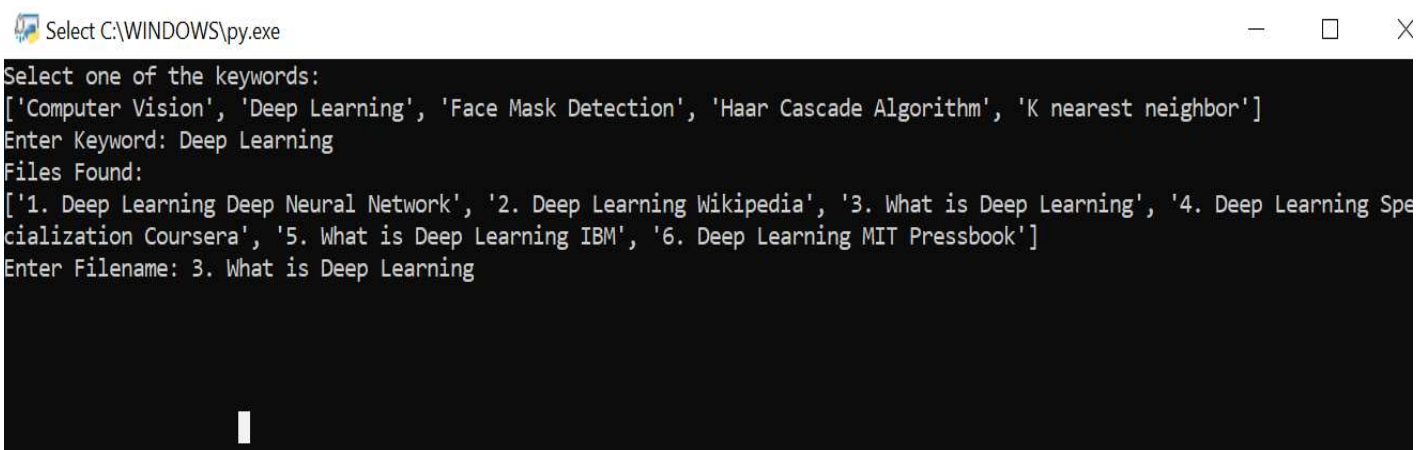
Figure 3- Storing the file with the keyword, filename and path in MongoDB

II. Repository Maintenance

Then **Python was used for the Repository Maintenance** step. Packages like os, sys and pymongo were installed. The connection was set up with the database. The keywords were listed and the desired search keyword was retrieved from the user. Based on the keyword, the database was searched and the corresponding list of files are returned. After the user types the desired file name, the file is opened.

```
C: > Users > aadil > OneDrive > Desktop > lab2 > lab3.py
1  import os
2  import sys
3
4  import pymongo
5
6  client = pymongo.MongoClient("mongodb://localhost:27017/")
7  database = client["ir"]
8  collection = database["webdata"]
9
10 print("Select one of the keywords:")
11 print(['Computer Vision', 'Deep Learning',
12        'Face Mask Detection', 'Haar Cascade Algorithm', 'K nearest neighbor'])
13 keyword = input("Enter Keyword: ")
14
15 query = {"keyword": keyword}
16 try:
17     document = list(collection.find(query))[0]
18 except IndexError:
19     print("Keyword not found")
20     sys.exit()
21
22 print("Files Found: ")
23 print(list(document['files'].keys()))
24
25 filename = input('Enter Filename: ')
26 try:
27     os.startfile(document['files'][filename])
28 except KeyError:
29     print("File not found")
30     sys.exit()
31
```

Figure 4 - Python Program for Repository Maintenance



```
Select C:\WINDOWS\py.exe
Select one of the keywords:
['Computer Vision', 'Deep Learning', 'Face Mask Detection', 'Haar Cascade Algorithm', 'K nearest neighbor']
Enter Keyword: Deep Learning
Files Found:
['1. Deep Learning Deep Neural Network', '2. Deep Learning Wikipedia', '3. What is Deep Learning', '4. Deep Learning Specialization Coursera', '5. What is Deep Learning IBM', '6. Deep Learning MIT Pressbook']
Enter Filename: 3. What is Deep Learning
```

Figure 5- Interactive Python Terminal

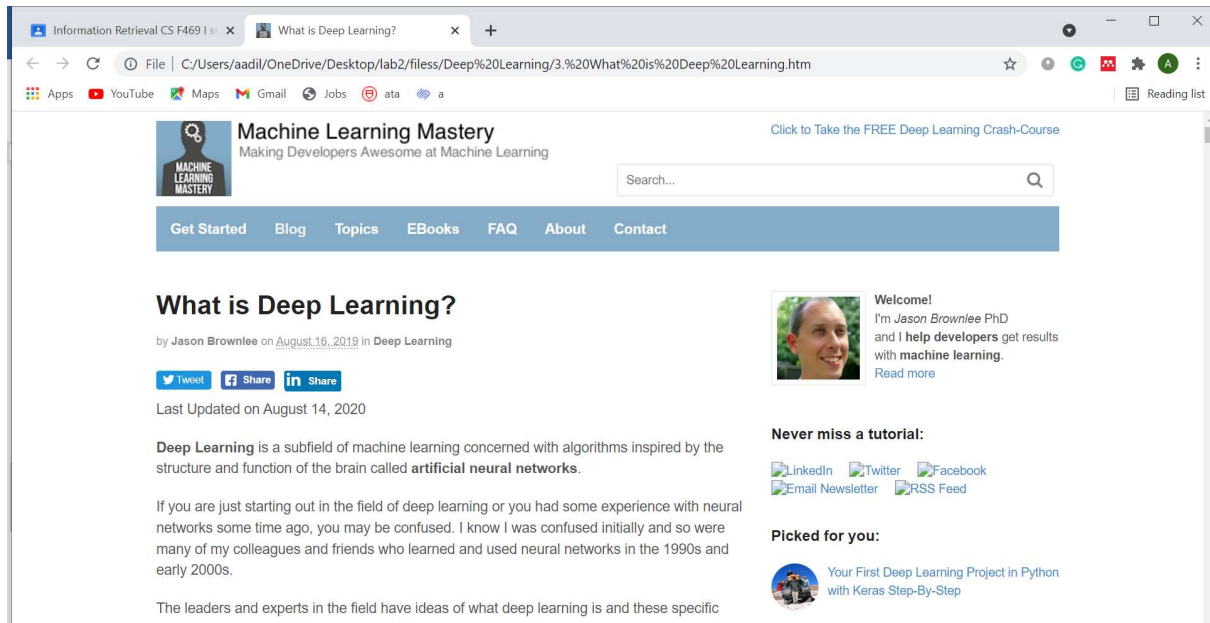


Figure 6 - Displays the desired webpage

Based on the selection from the user, the website is opened.