



VIT[®]

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

Register Number: 18BCE0745

Name: Gourishetty Sreemanth

Code:

```
# -*- coding: utf-8 -*-
```

```
''''
```

```
Created on Thu Jul 16 17:07:37 2020
```

```
@author: Sreemanth
```

```
''''
```

```
import re
```

```
fopen1 = open('Artificaial intelligence.txt', 'r')
```

```
for line1 in fopen1:
```

```
    print(line1)
```

```
fopen2 = open('machine learning.txt', 'r')
```

```
for line2 in fopen2:
```

```
    print(line2)
```

```
fopen1 = open('Artificaial intelligence.txt', 'r')
```

```
data1 = fopen1.read()
words1 = data1.split()
print('Number of words in text file1 :', len(words1))
```

```
fopen2 = open('machine learning.txt', 'r')
```

```
data2 = fopen2.read()
words2 = data2.split()
print('Number of words in text file1 :', len(words2))
```

```
fopen = open('Artificiaial intelligence.txt','r')
text_string = fopen.read().lower()
match_pattern = re.findall(r'\b[a-z]{3,15}\b', text_string)
frequency = {}
for word in match_pattern:
    count = frequency.get(word,0)
    frequency[word] = count + 1
frequency_list = frequency.keys()
```

```
for words in frequency_list:
    print (words, frequency[words])
```

```
print("#####COMMON WORDS -4")
```

```
#Group in a list the words common for two text files and show their total count
```

```
f1 = open("machine learning.txt").readlines()
f2 = open("Artificiaial intelligence.txt").readlines()
if len(f1) != 0 | len(f2) != 0:
    uniq1 = set(words for line in f1 for words in line.strip().split())
```

```
uniq2 = set(wordss for lines in f2 for wordss in lines.strip().split())
```

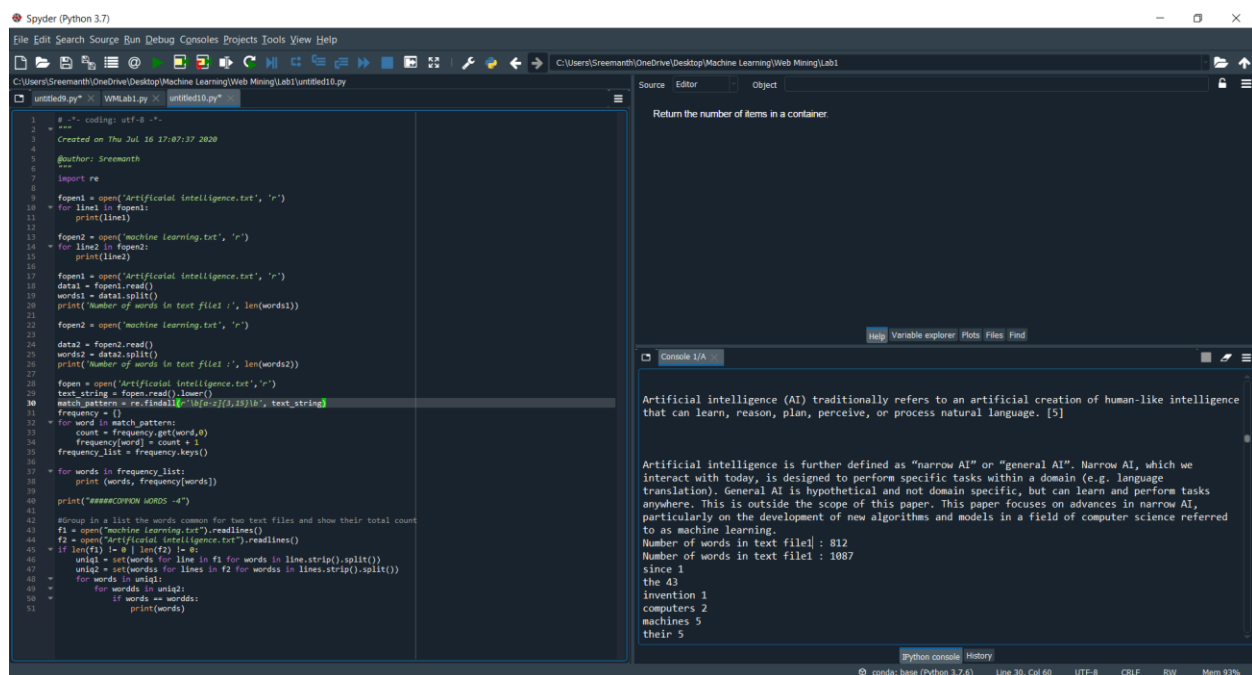
```
for words in uniq1:
```

```
    for wordds in uniq2:
```

```
        if words == wordds:
```

```
            print(words)
```

Output:



```
1 # -*- coding: utf-8 -*-
2 """
3 Created on Thu Jul 16 17:07:37 2020
4
5 @author: Sreemuth
6 """
7 import re
8
9 fopen1 = open('Artificial Intelligence.txt', 'r')
10 for line1 in fopen1:
11     print(line1)
12
13 fopen2 = open('machine learning.txt', 'r')
14 for line2 in fopen2:
15     print(line2)
16
17 fopen1 = open('Artificial Intelligence.txt', 'r')
18 data1 = fopen1.read()
19 words1 = data1.split()
20 print('Number of words in text file1:', len(words1))
21
22 fopen2 = open('machine learning.txt', 'r')
23
24 data2 = fopen2.read()
25 words2 = data2.split()
26 print('Number of words in text file2:', len(words2))
27
28 fopen = open('Artificial Intelligence.txt', 'r')
29 text_string = fopen.read().lower()
30 match_pattern = re.findall(r'\b[a-z]{3,10}\b', text_string)
31 frequency = {}
32 for word in match_pattern:
33     count = frequency.get(word, 0)
34     frequency[word] = count + 1
35 frequency_list = frequency.keys()
36
37 for words in frequency_list:
38     print(words, frequency[words])
39
40 print("=====COMMON WORDS=====")
41
42 #Group in a list the words common for two text files and show their total count
43 f1 = open('machine learning.txt').readlines()
44 f2 = open('Artificial Intelligence.txt').readlines()
45 if len(f1) != 0 and len(f2) != 0:
46     uniq1 = set(words for line in f1 for words in line.strip().split())
47     uniq2 = set(words for lines in f2 for words in lines.strip().split())
48     for words in uniq1:
49         for wordds in uniq2:
50             if words == wordds:
51                 print(words)
```

Return the number of items in a container.

Help: Variable explorer | Plots | Files | Find

Console I/O

Artificial intelligence (AI) traditionally refers to an artificial creation of human-like intelligence that can learn, reason, plan, perceive, or process natural language. [5]

Artificial intelligence is further defined as "narrow AI" or "general AI". Narrow AI, which we interact with today, is designed to perform specific tasks within a domain (e.g. language translation). General AI is hypothetical and not domain specific, but can learn and perform tasks anywhere. This is outside the scope of this paper. This paper focuses on advances in narrow AI, particularly on the development of new algorithms and models in a field of computer science referred to as machine learning.

Number of words in text file1: 812
Number of words in text file2: 1087

since 1
the 43
invention 1
computers 2
machines 5
their 5

Python console | History

conda: base (Python 3.7.4) | Line 30, Col 60 | UTF-8 | CRLF | RW | Mem 93%

Spyder (Python 3.7)

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\Sreemanth\OneDrive\Desktop\Machine Learning\Web Mining\Lab1\untitled10.py

untitled9.py* untitled10.py* untitled10.py*

```
1 # -*- coding: utf-8 -*-
2 """
3 Created on Thu Jul 16 17:07:37 2020
4
5 @author: Sreemanth
6 """
7 import re
8
9 fopen1 = open('Artificial Intelligence.txt', 'r')
10 for line1 in fopen1:
11     print(line1)
12
13 fopen2 = open('machine learning.txt', 'r')
14 for line2 in fopen2:
15     print(line2)
16
17 fopen1 = open('Artificial Intelligence.txt', 'r')
18 data1 = fopen1.read()
19 words1 = data1.split()
20 print('Number of words in text file1 :', len(words1))
21
22 fopen2 = open('machine learning.txt', 'r')
23
24 data2 = fopen2.read()
25 words2 = data2.split()
26 print('Number of words in text file2 :', len(words2))
27
28 fopen = open('Artificial Intelligence.txt', 'r')
29 text_string = fopen.read().lower()
30 match_pattern = re.findall(r'\b[a-z]{3,10}\b', text_string)
31 frequency = {}
32 for word in match_pattern:
33     count = frequency.get(word,0)
34     frequency[word] = count + 1
35 frequency_list = frequency.keys()
36
37 for words in frequency_list:
38     print(words, frequency[words])
39
40 print("#####COMMON WORDS -4")
41
42 #Group in a list the words common for two text files and show their total count
43 f1 = open('machine learning.txt').readlines()
44 f2 = open('Artificial Intelligence.txt').readlines()
45 if len(f1) != len(f2):
46     unig1 = set(words for line in f1 for words in line.strip().split())
47     unig2 = set(words for line in f2 for words in line.strip().split())
48     for words in unig1:
49         for words in unig2:
50             if words == words:
51                 print(words)
```

Return the number of items in a container.

Help Variable explorer Plots Files Find

Console I/O

```
efficient 1
processors 1
memory 1
addition 1
mistakes 1
adapt 1
environment 1
#####COMMON WORDS -4
many
recognition
like
learning,
is
science
used
trying
natural
a
and
making
various
```

Python console History

conda: base (Python 3.7.6) Line 30, Col 60 UTF-8 CRLF RW Mem 92%