Name: Venkata Krishnarjun Vuppala **Semester**: 6
SRN: PES2UG19CS451 **Section**: G
Subject: Algorithms for Intelligence Web and Information Retrieval

Assignment 1

1.Write a program to find word frequency in a given text file.

Code:

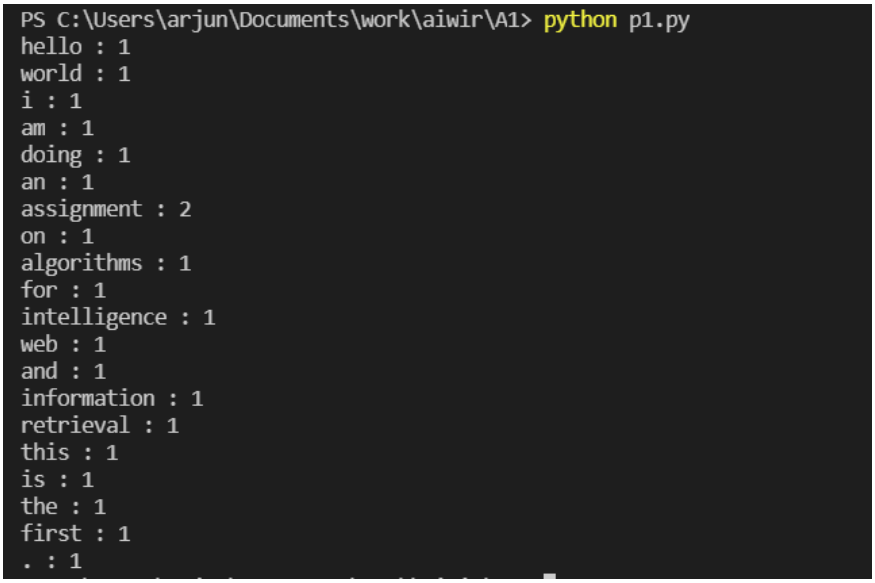
```
data = open("file.txt","r")
word_frequency = {}
for line in data:
    line = line.strip()
    line = line.lower()
    words = line.split(" ")
    for word in words:
        if word in word_frequency:
            word_frequency[word] += 1
        else:
            word_frequency[word] = 1
for key in word_frequency:
    print(key, ":", word_frequency[key])
```

File:

file.txt

```
1  Hello world
2  I am doing an assignment on Algorithms for Intelligence Web and Information Retrieval
3  This is the first assignment .
```

Output:



2. Write a program using array and linked list to -

- Store words
- Insert new word
- Search for a given word and return its index position
- Access element in the first, last and a particular position
- Compare both the data structures and analyze which approach is efficient.

Array:

Input:

p2.py

```python
import time

begin = time.time()

words = ['apple', 'banana', 'pineapple', 'tomato', 'chicken']

print("Array of words = ",words)

searchword = "apple"
for i in range(len(words)):
    if searchword == words[i]:
        print("Word found at index = ", i)

print("The first word in the array is: ", words[0])
print("The last word in the array is: ", words[-1])
end = time.time()

print(f"Total runtime of the program is {end - begin}")
```

Output:

```
PS C:\Users\arjun\Documents\work\aiwir\A1> python p2.py
Array of words =  ['apple', 'banana', 'pineapple', 'tomato', 'chicken']
Word found at index =  0
The first word in the array is:  apple
The last word in the array is:  chicken
Total runtime of the program is 0.001773834228515625
```

Linked List:

Code:

p3.py

p3.py > LinkedList > insAtBeg

```
4   class node:
5       def __init__(self,data):
6           self.data = data
7           self.next = None
8
9   class LinkedList:
10      def __init__(self):
11          self.head = None
12
13      def insAtBeg(self,new_data):
14          newNode = node(new_data)
15          newNode.next = self.head
16          self.head = newNode
17
18      def insAfter(self,prevNode,new_data):
19          if prevNode is None:
20              print("Invalid!!")
21              return
22          newNode = node(new_data)
23          newNode.next = prevNode.next
24          prevNode.next = newNode
25
26      def insAtEnd(self, new_data):
27          newNode = node(new_data)
28          if self.head is None:
29              self.head = newNode
30              return
31          last = self.head
32          while (last.next):
33              last = last.next
34          last.next = newNode
```

p3.py

p3.py > LinkedList > insAtBeg

```
    def search(self, key):
        curr = self.head
        while curr:
            if curr.data == key:
                return True
            curr = curr.next
        return False

    def lastelement(self):
        last = self.head
        while last.next:
            last = last.next
        print("The last word in the linked list is ",last.data)

    def printList(self):
        temp = self.head
        while (temp):
            print(str(temp.data) + " ", end="")
            temp = temp.next


llist = LinkedList()
llist.insAtEnd("Apple")
llist.insAtBeg("Mosambi")
llist.insAtBeg("Orange")
llist.insAtEnd("Tomato")
llist.insAfter(llist.head.next, "Mango")
print('linked list:')
llist.printList()

print()
item_to_find = "Orange"
if llist.search(item_to_find):
    print(str(item_to_find) + " is found")
else:
    print(str(item_to_find) + " is not found")


item_to_find = "banana"
if llist.search(item_to_find):
    print(str(item_to_find) + " is found")
else:
    print(str(item_to_find) + " is not found")


print("The first word in the linked list is",llist.head.data)
llist.lastelement()
end = time.time()
print(f"Total runtime of the program is {end - begin}")
```

Output:

```
PS C:\Users\arjun\Documents\work\aiwir\A1> python p3.py
linked list:
Orange Mosambi Mango Apple Tomato
Orange is found
banana is not found
The first word in the linked list is Orange
The last word in the linked list is  Tomato
Total runtime of the program is 0.0009965896606445312
PS C:\Users\arjun\Documents\work\aiwir\A1>
```