BIOINFORMATICS

LAB ASSIGNMENT 2

Longest Comment Subsequence

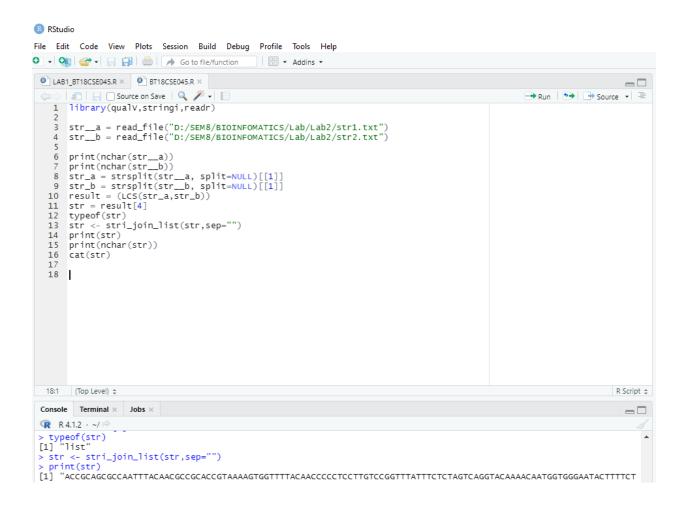
Submitted by

Nitesh Yadav BT18CSE045

PYTHON

```
🕏 subseq.py > ...
  1 from difflib import SequenceMatcher
     str a = open("/home/nik/cp/str1.txt")
     str_b = open("/home/nik/cp/str2.txt")
     str a = str a.read()
     str b = str b.read()
     def lcs(s1, s2):
        matrix = [["" for x in range(len(s2))] for x in range(len(s1))]
         for i in range(len(s1)):
            for j in range(len(s2)):
                if s1[i] == s2[j]:
                   if i == 0 or j == 0:
                       matrix[i][j] = s1[i]
                   else:
                       matrix[i][j] = matrix[i-1][j-1] + s1[i]
                   matrix[i][j] = max(matrix[i-1][j], matrix[i][j-1], key=len)
                   cs = matrix[-1][-1]
 19
         return len(cs), cs
     print(lcs(str a, str b))
PROBLEMS OUTPUT TERMINAL
```

Advantages	Disadvantages
 Problem-solving skills are enhanced Learning dynamic programming 	 Fully Coding the approach is tough Learning the approach and dynamic programming is time-consuming



Advantages	Disadvantages
 Vast Variety of Libraries Learned about File I/O operations in R Learned about the modules & functions such qualV, LCS 	 Finding the exact packages takes is time-consuming in the R documentation Limited resources available online