Assignments

Homework 6

Objective
Use Python to create a program using lists in the DNA domain.

Problem domain
For two strings s1 and s2 of equal length, the p-distance between them, denoted dp(s1,s2), is the proportion of corresponding symbols that differ between s1 and s2.

For a general matrix distance function d on n taxa s1,s2,…,sn (taxa are often represented by genetic strings), we may encode the distances between pairs of taxa via a distance matrix D in which Di,j=d(si,sj).

Given: A collection of n (n≤10) DNA strings s1,…,sn of equal length (at most 1 kbp). Strings are given
in FASTA format.

Return: The matrix D corresponding to the p-distance dp on the given strings. As always, note that your answer is allowed an absolute error of 0.001.

```
Sample
Dataset
          [
           ['T','T','T','C','C','A','T','T','T','A'],#list1
           ['G','A','T','T','C','A','T','T','T','C'],#list2
           ['T','T','T','C','C','A','T','T','T','T'],#list3
           ['G','T','T','C','C','A','T','T','T','A']#list4
          ]

          Sample Output
          0.00000 0.40000 0.10000 0.10000 #firstrow
          0.40000 0.00000 0.40000 0.30000
          0.10000 0.40000 0.00000 0.20000
          0.10000 0.30000 0.20000 0.00000
```

Tips:
Compare list1 to list2, list3, and list4 to get the p distance for the first row. Do the same for the other lists. Compare list2 to list1, list3, and list4. Etc.

Use the get_p_distance function in the get_p_distance_matrix function.

Prerequisites
Install Python
GitHub account and repository

Install and configure Visual Studio Code

Write Code

In Visual Studio Code, find the /src/homework/i_dictionaries_sets folder.

a. In the dictionary.py file, write the value return functions:
get_p_distance with list parameter list1 and list2 (see get p distance above for function code)
get_p_distance_matrix with list parameter list1 (see general matric function above for function code)
Use the get_p_distance function to get the distance between two lists, save the result to p_distance_matrix[i][j].

b. Write the Tests for the functions (see next section)

Write Unit Test

In Visual Studio Code, find the /tests/homework/i_dictionaries_sets folder.

a. In the file test_dictionaries_and_sets.py add the following code:
import unittest
from src.homework.i_dictionaries_and_sets import get_p_distance
from src.homework.i_dictionaries_and_sets import get_p_distance

class Test_Config(unittest.TestCase):

b. After the line that begins with class write a test case function test_p_distance
Test that get_p_distance with parameter values
`['T','T','T','C','C','A','T','T','T','A']` and
`['G','A','T','T','C','A','T','T','T','C']` that returns .4 .

c. Test case function test_get_p_distance_matrix
Test that get_p_distance matrix with parameter value
[
['T','T','T','C','C','A','T','T','T','A'],
['G','A','T','T','C','A','T','T','T','C'],
['T','T','T','C','C','A','T','T','T','T'],
['G','T','T','C','C','A','T','T','T','A']
]

returns
[
[0.0, 0.4, 0.1, 0.1],
[0.4, 0.0, 0.4, 0.3],
[0.1, 0.4, 0.0, 0.2],
[0.1, 0.3, 0.2, 0.0]
]

Run the Unit Tests

In Visual Studio Code, find the /tests/homework/i_dictionaries_sets folder

a. From the source code root folder, find the run_tests.py file.
Replace: `from tests.homework.h_strings import test_strings`
with
`from tests.homework.i_dictionaries_and_sets import tests_dictionaries_and_sets`
Verify that line has the following statement :
Replace tests_strings with tests_dictionaries_and_sets
`suite = unittest.TestLoader().loadTestsFromModule(tests_dictionaries_and_sets)`
b. Click on the play button to run the test case.
c. Make sure the test results return ok for the test cases (Fix the code if it fails).

Create and Run the Main Program

In Visual Studio Code, find the /src/homework/I_dictionaries _sets folder find the main.py file. write code to create the following menu.

1-Get p distance matrix
2-Exit

The program runs until the user chooses option 2

Option 1 prompt the user for a list, call the get_p_distance_matrix function and display the result.

Upload the Changes to GitHub

a. In Visual Studio Code, click on the Source Control icon .
b. Select only the files pertaining to this assignment.
c. Click on the + to stage the changes.
d. Click on the check mark to commit the changes.
e. From the menu select the …, from the menu select Push.

Submit the Assignment for Grading in Blackboard

Make sure to add your GitHub user name to the Comment edit box.

Assignment 6

What is a list? Draw the list memory diagram for the following code.

list1 = [4,8,10]