

CS 412/512 - Homework #1

Due To

8 March 2022 Tuesday, by 11.59 pm.

To Do

Write appropriate functions in (Java or Python) language for the following problems:

Frequent Words Problem (1B):

Find the most frequent k-mers in a string.

Input: A String Text and an integer k.

Output: All most frequent k-mers in Text.

Reverse Complement Problem (1C):

Find the reverse complement of a DNA string.

Input: A DNA string Pattern

Output: Pattern, the reverse complement of Pattern.

Pattern Matching Problem (1D):

Find all occurrences of a pattern in a string.

Input: String pattern and Genome

Output: All starting positions in Genome where Patter appears as a substring

Clump Finding Problem (1E):

Find patterns forming clumps in a string.

Input: A string Genome, and integers k, L, and t.

Output: All distinct k-mers forming (L, t)-clumps in Genome.

Rules

- You are not allowed to use any external libraries and/or functions. Everything must be implemented by you from scratch.
- Submit ".py" or ".java" file. Any other type of file (such as .ipynb and etc) will be disregarded.
- It is your responsibility to make sure that your code runs without any type of errors. If your code does not compile or run, you may only get partial credits.
- Any type of plagiarism will not be tolerated. Your submitted codes will be compared with other submissions and also the codes available on internet and violations will have a penalty of -100 points. (In case of copying from another student both parties will get -100)
- Please submit your code using the specified file name as "HW#_YourName_YourSurname.(py/java)" for instance for homework 1, it would be "HW1_Jane_Johnson.py" or "HW1_James_Johnson.java"