CSCE 300 Assignment 2

Due: February 21, 2019

In this assignment, you will develop three algorithm from the text book.

1- Write a Python code for **ApproximatePatternCount** algorithm.
2- Write a Python code for **FrequentWordsWithMRC** algorithm.
3- Find and report potential DnaA boxes for Vibrio Cholerae by running **FrequentWordsWithMRC** (OriC of Vibrio Cholerae) – Data can be found on BeachBoard.
4- Find and report potential DnaA boxes for E. Coli by running **FrequentWordsWithMRC** (OriC of E.Coli) – Data can found on BeachBoard.

After finishing the code you need to submit the following:

1- Your source code (freqword.py).
2- A text file that contains your program output for the given input (output.txt).
3- A Readme file (Readme.txt) that contains the following information
   a. The program authors name
   b. Your program result (you can copy and paste from the terminal)
   c. How do you interpret the generated result by the program

The submission format and rules (Some coding guideline):

1- For this assignment, send only one .py file.
2- On top of your code, you need to have proper description header (comments).
3- In the code if needed add inline description as comments.
4- All function requires comments before the function name describing the functionality of the function.

Resources:

- A review paper about OriC: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4285127/
- A video about DNA replication in 3D: https://www.youtube.com/watch?v=TNKWgcFPHqw