September 17, 2023

**CS474 – Bioinformatics**

**Assignment 1**

From GenBank database located at <https://ftp.ncbi.nlm.nih.gov/genbank/> download a file with digits in their names match last three digits of your registration number. After downloading the file, analyze each record using **BioPython** as follows:

1. Print each record ID and total number of records in the file.
2. Compute the GC fraction of the sequence in each record (*Note: the gc\_fraction*( ) *function is case sensitive*).
3. If the sequence in each record is a coding strand, then compute the template strand of the sequence.
4. Transcribe the sequence in each record into mRNA.
5. Translate each mRNA into proteins and save them in a file in GenBank format along with other important features.

**Report is due on or before Sunday, September 24, at 11:59PM on MS Teams.**