Syllabus
BINF-734 Advanced Programming for Bioinformatics
CDS-486 Topics in Computational and Data Sciences

1   Information

Title: BINF-734 Advanced Programming for Bioinformatics
   CDS-486 Topics in Computational and Data Sciences.
Instructor: Dr. Jason M. Kinser
Location: Fairfax Campus (Room TBD)
Time: Tuesday 7:20pm - 10:00pm
Text: Documents provided by the instructor and "Python for Bioinformatic" J. Kinser, Jones & Bartlett, 2008.
Prerequisite: Experience with any procedural programming language (C, C++, Java, Fortran, etc.)

2   Purpose

The purpose of this class is for the student to gain an understanding of how to create a variety of programs to perform different bioinformatics tasks. The course will include descriptions of algorithms and methods by which these can be implemented into computer codes. This course will also attempt to replicate some recent bioinformatics publications.

3   Prerequisites

Previous programming experience is not required but is an advantage. Students with experience in Java, C/C++, C# or other procedural languages will be able to draw on this experience.

4   Course Pedagogy

Grading in this course is based upon:
BINF-734 Students

- 70% of the grade is based upon the programming assignments, and

- 30% of the grade is based upon a final project.

CDS-486 Students

- 100% of the grade is based upon the programming assignments and attendance to the oral presentations.

4.1 Schedule

- Week 1: Intro to the course and Python.

- Week 2: Program to read Genbank files.

- Week 3: Sequence Alignments. Python arrays and numpy.

- Week 4: Dynamic Programming.

- Week 5: Multiple Alignments.

- Week 6: Clustering (Genetic algorithms and k-means)

- Week 7: Principle Component Analysis

- Week 8: Text Mining

- Week 9: Phylogenetic Trees

- Weeks 10-13. Topics will be selected based on student Master's or PhD research interests. Typical topics are:

  - Numerical and Image Representations of sequences
  - Self organizing maps
  - Basic Image Analysis
  - Skew Analysis
  - n-grams
  - Decision Trees
  - Normalization of gene expression arrays
  - Introduction to MySQL
  - New topics in recent publications

- Week 14: Semester Project Presentations

4.2 Programming Assignments

Student's will be required to write small programs for bioinformatics applications. These assignments are given out weekly excepting the first week, the last week, and on occasions where some topics last two weeks. Historically, there are 9 - 11 assignments within a semester.

Each programming assignment will consist of the following:

- Codes generated by the student,

- A small 1 or 2 page report that includes instructions on how to run the student's program and the results, and

- Data files generated by the student's program.

5 The Final Project

The final project constitutes 30% of the total grade for BINF-734 students.

The final project consists of the following steps:

1. Select a recent publication in the field of bioinformatics,

2. Obtain approval for the use of this publication, (you may change later but approval is required),

3. Obtain or generate data for this problem,

4. Identify the algorithms that are needed to replicate this paper,

5. Obtain or write the necessary programs, (In some cases, authors provide Python scripts. These are not allowed to be used as the project.)

6. Replicate a portion of the published work,

   - See the instructor about the amount of work that needs to be done,
   - In almost all cases the student will replicate only a small part of the published work,

7. Create an oral presentation.

8. Present the work on the last day of classes. (In some years it is necessary to also present on the day when the final would have been given.)

There are multiple project deadlines throughout the semester:

- Week 3. Submit a chosen publication for instructor approval. You may submit more than one, but indicate your level of choice,

- Week 6. Submit a list of algorithms that your project will require. This can be as simple as a list such as,

  - Simulated Annealing,
  - Neural Networks,

- Week 8. Data collected or generated.

- Week 12. Working programs.

- Week 14. Presentations