THE UNIVERSITY OF TEXAS AT EL PASO
COLLEGE OF SCIENCE
Bioinformatics Program

Course #: BINF 5112 (CRN 25390)
Course Title: Computer Science Seminar for Bioinformatics
Credit Hrs: 1
Term: Spring 2014
Course Meetings & Location: Wed. 12:30 – 1:20 p.m.; Education Building 108
Prerequisite Courses: None
Instructor: Ming-Ying Leung
Office Location: Bell Hall 225
Contact Info: Phone # 747-6836
E-mail mleung@utep.edu
Fax # 747-6502
Website http://www.bioinformatics.utep.edu/mleung
Office Hours: Dr. Leung will post her office hours for each week on her website and
outside her office, or by appointment.

Textbook(s), Materials:


Course Objectives (Learning Outcomes):

Machine Learning studies the development of programs that can improve in
the performance of a task with experience. For many difficult problems,
solutions based on machine learning outperform all other solutions proposed
to date. Examples of these problems include speech recognition,
classification of objects in images, weather prediction, fraud detection, robot
navigation, and many others.

In this course we will discuss several of the most commonly used machine learning algorithms and their application to problems in bioinformatics. Most
meetings will consist of a presentation of either a paper from the scientific
literature or a chapter from a textbook, followed by a discussion session.

Assessment of Course Objectives:

During in-class presentations and discussions, I shall assess your ability:

- To understand basic machine learning algorithms commonly used
- To use standard machine learning and data mining software tools for solving bioinformatics problems.
Course Activities/Assignments: Paper Presentations:

Course Schedule:

1) Introduction
   a) What is Machine Learning?
   b) Learning algorithms

2) Learning Algorithms
   a) Neural Networks
      i) Feed forward neural networks
   b) Decision trees
      i) ID3
      ii) C4.5
   c) Graphical Models
      i) Hidden Markov models
      ii) Bayes nets
      iii) Conditional random fields
   d) Kernel Methods
      i) Support Vector Machines
   e) Evolutionary Algorithms
      i) Genetic algorithms
   f) Instance-based learning
      i) k-nearest neighbors
      ii) Locally-weighted regression
   g) Probabilistic Methods
      i) Expectation maximization
      ii) Fisher’s linear discriminant
   h) Ensembles of classifiers
      i) Boosting
      ii) Bagging
      iii) Randomization
      iv) Stacking
      v) Error-correcting output coding

3) Applications
   a) Folding prediction
      i) RNA folding
      ii) DNA folding
      iii) Protein folding.
   b) Mining the biological literature
   c) Gene regulatory networks
   d) Finding replication origins in DNA
   e) Gene finding

4) Tools
   a) WEKA
   b) Matlab

Grading Policy: Class participation and presentations 100%

Make-up Policy: IF you notify the instructor BEFORE the internship evaluation is complete, it may be possible to submit another internship proposal for approval.

Attendance Policy: Students missing more than two class meeting will receive an “F” for the course. Two tardies will count as one absence. A tardy will be recorded each time a student shows up five minutes after the start of class.


Civility Statement: Acceptable working relationship with the named internship supervisor and other personnel in the environment is expected.
Disability Statement: If a student has or suspects she/he has a disability and needs an accommodation, he/she should contact the Disabled Student Services Office (DSSO) at 747-5148 or at <dss@utep.edu> or go to Room 106 Union East Building. The student is responsible for presenting to the instructor any DSS accommodation letters and instructions.

Military Statement: If you are a military student with the potential of being called to military service and/or training during the course of the semester, you are encouraged to contact as soon as possible.