CS 4361/5361 Machine Learning
Spring 2023

Instructor:

Olac Fuentes
Email and MS Teams: ofuentes@utep.edu
Web: www.cs.utep.edu/ofuentes
Office hours: Mondays and Wednesdays 1:00 - 2:30 p.m. or by appointment in CCSB 3.0412 or MS Teams. Feel free to contact on MS Teams at other times if available.

Teaching Assistant (TA):

Md. Mahmudulla Hassan
Email and MS Teams: mhassan@miners.utep.edu
Office hours: TBA

Meeting times and place: M-W 3:00:4:20 p.m. in CCSB G.0208

Introduction:
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 study several of the most commonly used machine learning algorithms and their application to problems in several areas of interest. We will also discuss current research issues in Machine Learning and each student will do a research project related to a problem of his/her interest.

Course Contents:
1) Introduction
   a) Machine Learning – what, why and how.
   b) Machine Learning tasks: classification, regression, and reinforcement learning
   c) Review – Python and arrays
2) Performance evaluation
   a) Metrics
   b) Cross-validation
      i) Training and test sets
      ii) N-fold cross validation
      iii) Leave-one-out cross validation
3) Data preprocessing
   a) Normalization
   b) Standardization
   c) Principal component analysis
   d) One-hot encoding
4) Instance-based learning
   a) K-nearest neighbors
   b) Locally-weighted regression
5) Probabilistic algorithms
   a) Probability review
   b) The Naïve Bayes classifier

6) Tree-based algorithms
    a) Decision trees
    b) Regression trees
7) Ensembles of models
    a) Bagging and boosting
    b) Random forests
    c) Gradient boosting
8) Linear regression
9) Gradient descent and logistic regression
10) Feed-forward dense neural networks
11) Support vector machines
12) Convolutional neural networks
13) Unsupervised Learning
    a) Principal component analysis
    b) k-means
14) Semi-supervised learning
15) Self-supervised learning

Pre-requisites:
CS 2302 Data Structures, MATH 3323 Matrix Algebra, STAT 3320 Probability and Statistics, and ability to program in Python.

Grading

CS 4361:
Labs 25%
Class participation, homework, quizzes, and in-class exercises 15%
Partial exams (2) 50%
Final Project 10%

CS 5361:
Labs 25%
Class participation, homework, quizzes, and in-class exercises 15%
Partial exams (2) 40%
Final Project 20%

As reflected by the relative weights, the expectations for the final project are much higher for students taking CS 5361 or 6361.

Textbook:
There will be no textbook. We will use parts of online books and other materials.

Standards of Conduct and Academic Dishonesty:
You are expected to conduct yourself in a professional and courteous manner, as prescribed by the UTEP Standards of Conduct: https://admin.utep.edu/portals/68/Standards_of_Conduct_Booklet_5-11-15.pdf

Academic dishonesty includes but is not limited to cheating, plagiarism and collusion. Cheating may involve copying from or providing information to another student, possessing unauthorized materials during a test, or falsifying data (for example program outputs) in laboratory reports. Plagiarism occurs when someone represents the work or ideas of another person as his/her own. Collusion involves collaborating with another person to commit an academically dishonest act.

Professors are required to report academic dishonesty and any other violation of the Standards of Conduct to the Dean of Students.