

# CSci 574 Artificial Intelligence (Machine Learning and Data Analysis)

Course Syllabus

Fall 2017

## Instructor

Derek Harter, Ph.D., Associate Professor  
Department of Computer Science  
Texas A&M University - Commerce  
Office: SCI 355  
Office Hours: T, W, Th, 1 - 3 pm  
E-mail: Derek.Harter@tamuc.edu

## Class Meetings

Lectures and course materials will be distributed through our University's eCollege online course system.

01W 83006 Web Based Class

## Course Description

This semester's AI course will focus on the sub field of machine learning with some work in the analysis of big data. We will look at classic unsupervised and supervised learning methods, used in the field to classify data, cluster it and find optimizations. This will include looking at k-means and hierarchical clustering, self-organizing maps, linear regression, decision trees, optimization techniques such as genetic algorithms, etc. We will cover ways of getting hold of interesting datasets, ideas on how to collect data from users, and many different ways to analyze and understand the data once you've found it.

## Requirements and Objectives

- Explore downloading and mining real web data sets.
- Learn about unsupervised methods for grouping and visualization.
- Program optimization algorithms to search for optimal solutions using hill climbing, GA's, etc.
- Become familiar with some advanced classification techniques of the kernel methods family of algorithms.

## Companion Textbooks / Readings

- [CI] Programming Collective Intelligence by Toby Segaran, O'Reilly Media, 2007, ISBN: 0596529325
- [ML] Marsland. (2009). Machine Learning: An Algorithmic Perspective. By Marsland, CRC Press, 2009.
- [PY] How to Think Like a Computer Scientist: Learning with Python 2ed by Jeffrey Elkner, Allen B. Downey and Chris Meyers (Open Book Project) <http://www.greenteapress.com/thinkpython/>

## Prerequisites

Background in basic programming (Undergraduate equivalent of CSci 151/152) and Data Structures (Undergraduate equivalent of CSci 270) or equivalent knowledge will be assumed for this course.

## Evaluation (Tentative)

Your grade for the course will be based on the following (approximate) percentages:

Two Exams (Midterm and Final)	50% (25% each)
Labs / Programming Assignments (appx. 4-5)	45%
Class Participation	5%

Letter grades will be assigned according to the following scale:

Final Average (%)	Letter Grade
90 - 100	A
80 - 89	B
70 - 79	C
60 - 69	D
Below 60	F

## **Student's with Disabilities**

The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you have a disability requiring an accommodation, please contact: Office of Student Disability Resources and Services, Texas A&M University-Commerce, Gee Library, Room 132, Phone (903) 886-5150, StudentDisabilityServices@tamuc.edu

## **Academic Ethics**

"All students enrolled at the University shall follow the tenets of common decency and acceptable behavior conducive to a positive learning environment." (See Student's Guide Handbook, Policies and Procedures, Conduct). Ethics also includes the issue of plagiarism, and copying code for programming/lab assignments is just as serious as any other type of plagiarism. If you are caught sharing or using other people's work in this class, you will receive a 0 grade and a warning on the first instance. A subsequent instance will result in receiving an F grade for the course, and possible disciplinary proceedings.

## **Attendance Policy**

Students are expected to follow all instructions and visit eCollege regularly many times weekly to complete the materials for this online course. If a student is unable to submit assignments by the due date for the assignment, they are expected to make alternative arrangements to assure that the assignment is turned in ON TIME, before the assignment is actually due. Any student wishing to withdraw from the course must do so officially as outlined

in the class schedule. THE INSTRUCTOR CANNOT DROP OR WITHDRAW ANY STUDENT.

## Course Requirement Deadlines

Credit will be given for ONLY those exam(s), program(s), and/or project(s) turned in no later than the deadline(s) as announced by the instructor of this class unless prior arrangement has been made with the instructor.

## Technology Requirements

This course is a web enhanced course, which means all assignments and handouts will be distributed and collected through our University's eCollege online course system.

- To fully participate in online courses you will need to use a current browser, such as Mozilla Firefox or Google Chrome.
- You will need regular access to a computer with a broadband internet connection. The minimum computer requirements are:
  - 512 MB of RAM, 1 GB or more preferred.
  - Broadband connection required by courses that are heavily video intensive.
  - Video display capable of high-color 16-bit display 1024 x 768 or higher resolution.
- You must have a:
  - Sound card, which is usually integrated into your desktop or laptop computer.
  - Speakers or headphones.

## Access and Navigation

This course will be facilitated using Pearson LearningStudio, the learning management system used by Texas A&M University-Commerce. To get started with the course, go to myLeo and from the top menu ribbon select eCollege. Then on the upper left side of the screen click on the My Courses tab. <http://www.tamuc.edu/myleo.aspx>

You will need your campus-wide ID (CWID) and password to log into the course. If you do not know your CWID or have forgotten your password, contact the Center for IT Excellence (CITE) at 903.468.6000 or [helpdesk@tamuc.edu](mailto:helpdesk@tamuc.edu).

Note: It is strongly recommended you perform a “Browser Test” prior to the start of your course. To launch a browser test login to Pearson LearningStudio, click on the My Courses tab, and then select the Browser Test link under Support Services.

Texas A&M University-Commerce provides students technical support for the use of Pearson LearningStudio. Technical assistance is available 24/7 (24 hours, 7 days a week). If you experience LearningStudio (eCollege) technical problems, contact the LearningStudio helpdesk at 1-866-656-5511 (toll free) or visit Pearson 24/7 Customer Support Site <http://247support.custhelp.com/>

Accessing Help from within Your Course: Click on the 'Tech Support' icon on the upper left side of the screen inside the course. Then you will be able to get assistance via online chat or by phone.

Note: Personal computer and internet connection problems do not excuse the requirement to complete all course work in a timely and satisfactory manner. Each student needs to have a backup method to deal with these inevitable problems. These methods might include the availability of a backup PC at home or work, the temporary use of a computer at a friend's home, the local library, office service companies, Starbucks, a TAMUC campus open computer lab, etc.

myLeo Support: Your myLeo email address is required to send and receive all student correspondence. Please email [helpdesk@tamuc.edu](mailto:helpdesk@tamuc.edu) or call us at 903-468-6000 with any questions about setting up your myLeo email account. You may also access information at myLeo. <https://leo.tamuc.edu>

## **Nondiscrimination Notice**

Texas A&M University-Commerce will comply in the classroom, and in on-line courses, with all federal and state laws prohibiting discrimination and related retaliation on the basis of race, color, religion, sex, national origin, disability, age, genetic information or veteran status. Further, an environment free from discrimination on the basis of sexual orientation, gender identity, or gender expression will be maintained.

## Campus Concealed Carry Statement

Texas Senate Bill - 11 (Government Code 411.2031, et al.) authorizes the carrying of a concealed handgun in Texas A&M University-Commerce buildings only by persons who have been issued and are in possession of a Texas License to Carry a Handgun. Qualified law enforcement officers or those who are otherwise authorized to carry a concealed handgun in the State of Texas are also permitted to do so. Pursuant to Penal Code (PC) 46.035 and A&M-Commerce Rule 34.06.02.R1, license holders may not carry a concealed handgun in restricted locations.

For a list of locations, please refer to the Carrying Concealed Handguns On Campus document and/or consult your event organizer.

Web url: <http://www.tamuc.edu/aboutUs/policiesProceduresStandardsStatements/rulesProcedures/34SafetyOfEmployeesAndStudents/34.06.02.R1.pdf>

Pursuant to PC 46.035, the open carrying of handguns is prohibited on all A&M- Commerce campuses. Report violations to the University Police Department at 903- 886-5868 or 9-1

## Course Schedule (Preliminary)

W	Date	Topic / Activity	Notebook	Test
1	8/28	Course Introduction, Set up Python	1a & 1b	
2	9/4	Introduction to Python Scientific Libraries	2a, 2b, 2c	
3	9/11	Linear Regression with One Variable		
4	9/18	Linear Regression with Multiple Variables		
5	9/25	Logistic Regression and Regularization		
6	10/2	K Nearest Neighbor		
7	10/9	Decision Trees and Ensembles		
8	10/16			Midterm
9	10/23	Support Vector Mechines (SVM) and Kernel Methods		
10	10/30	Unsupervised Learning and Dimensionality Reduction		
11	11/6	Principle Component Analysis (PCA)		
12	11/13	K-means and Hierarchical Clustering		
13	11/20	Anomaly Detection and Recommender Systems		
14	11/27	Large Scale Machine Learning		
15	12/4	Advice for Applying Machine Learning		
	12/11	Finals Week 12/12 - 12/16		Final

## Student Learning Outcomes

1. Develop familiarity with high-level Python scripting language.
2. Learn basics of fundamental machine learning techniques, such as regression, clustering, k-nearest neighbor, kernel methods, etc.
3. Learn basic distinction between supervised and unsupervised machine learning methods.
4. Learn basic machine learning training and testing techniques, including cross validation and data optimization.

Learning outcomes will be measured through mapping assignment and test questions to specific outcome items, as well as through exit surveys of student experiences with the outcome familiarity.