

CSCI 556.01W Data Analysis & Visualization

COURSE SYLLABUS: Fall 2021

(Revision date: August 16, 2021)

INSTRUCTOR INFORMATION

Instructor: Jinoh Kim, Ph.D. Office Location: CS/JOUR 217

Office Hours: Will be announced through course page

University Email Address: Jinoh.Kim@tamuc.edu

COURSE INFORMATION

Textbook(s): No textbook required (see an optional textbook below)

Software Required: Weka

Optional Texts and/or Materials: Data Mining: Practical Machine Learning Tools and Techniques, 4th Edition. ISBN-13: 978-0128042915, ISBN-10: 0128042915 (Optional)

Course Description

Big scientific data sets are growing exponentially both in size and complexity. Extracting meaningful information from these data requires not only programming skills, but also understanding the analysis work-flows, mathematical models and visualization tools that help to condense large amounts of information into a comprehensible story. We will introduce standard statistical data analysis and modeling methods such as correlation functions, linear regression, clustering, pattern extraction, classification, data mining, as well as Monte Carlo methods which are commonly used in creating simulations in the computational sciences. Different analysis and visualization packages popular in scientific modeling, analysis, and visualization will be introduced.

Student Learning Outcomes (Should be measurable; observable; use action verbs)

- 1. The student will gain detailed knowledge about the goal and techniques of the data analysis and visualization process.
- 2. The student will understand the steps in characterizing and understanding data.
- 3. The student will be able to build effective predictive models.
- 4. The student will be able to build models that rely on memorizing training data.
- 5. The student will be able to build models that have an explicit additive structure.
- 6. The student will be able to build models for data that has no labeled training data available: Unsupervised learning
- 7. The student will be able to use software applications for data analysis.

COURSE REQUIREMENTS

Prerequisites

None

Minimal Technical Skills Needed

- Basic knowledge of computer and operating systems
- Basic knowledge of data structure, algorithms, and statistics/probability

Instructional Methods

- Online lecture for data analysis & visualization concepts
- Will learn how to utilize an open source data analysis and visualization tool: WEKA (http://old-www.cms.waikato.ac.nz/~ml/weka/)

Student Responsibilities or Tips for Success in the Course

- On-time submission of assignments
- Should well be well (fully) prepared for exams
- IMPORTANT: This is an online-formatted course, and it is mandatorily required
 that the student reads announcements from the course page frequently (at least
 three times a week) and email messages from the instructor without any
 significant delay (less than 24 hours).

GRADING (Tentative)

Final grades in this course will be based on the following scale:

A = 90%-100%

B = 80%-89%

C = 70% - 79%

D = 60%-69%

F = 59% or Below

Weights of the assessments in the calculation of the final letter grade:

Components	Weight	Remarks
Assignments	40%	Homework assignments
Midterm exam	30%	Two exams and the highest score will be chosen
Final exam	30%	Cumulative

Late Policy (Assignments):

The deadline for the assignment can be extended with a 15% penalty per day, up to two days. No submission will be accepted 48 hours after the deadline.

Makeup Policy (Exams):

Makeup chances may be given to students under extreme circumstances only, such as hospitalization, serious injury, death in the family, etc, with prior consent <u>and</u> the supporting document (officially issued).

TECHNOLOGY REQUIREMENTS

LMS

All course sections offered by Texas A&M University-Commerce have a corresponding course shell in the myLeo Online Learning Management System (LMS). Below are technical requirements

LMS Requirements:

https://community.brightspace.com/s/article/Brightspace-Platform-Requirements

LMS Browser Support:

https://documentation.brightspace.com/EN/brightspace/requirements/all/browser_support.htm

YouSeeU Virtual Classroom Requirements:

The syllabus/schedule are subject to change.

https://support.youseeu.com/hc/en-us/articles/115007031107-Basic-System-Requirements

ACCESS AND NAVIGATION

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.

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.

COMMUNICATION AND SUPPORT

If you have any questions or are having difficulties with the course material, please contact your Instructor.

Technical Support

If you are having technical difficulty with any part of Brightspace, please contact Brightspace Technical Support at 1-877-325-7778. Other support options can be found here:

https://community.brightspace.com/support/s/contactsupport

COURSE AND UNIVERSITY PROCEDURES/POLICIES

Course Specific Procedures/Policies

N/A

Syllabus Change Policy

The syllabus is a guide. Circumstances and events, such as student progress, may make it necessary for the instructor to modify the syllabus during the semester. Any changes made to the syllabus will be announced in advance.

University Specific Procedures

N/A

The syllabus/schedule are subject to change.

Student Conduct

All students enrolled at the University shall follow the tenets of common decency and acceptable behavior conducive to a positive learning environment. The Code of Student Conduct is described in detail in the Student Guidebook.

 $\underline{http://www.tamuc.edu/Admissions/oneStopShop/undergraduateAdmissions/studentGuidebook.as}\\px$

Students should also consult the Rules of Netiquette for more information regarding how to interact with students in an online forum: https://www.britannica.com/topic/netiquette

TAMUC Attendance

For more information about the attendance policy please visit the <u>Attendance</u> webpage and <u>Procedure 13.99.99.R0.01</u>.

http://www.tamuc.edu/admissions/registrar/generalInformation/attendance.aspx

http://www.tamuc.edu/aboutUs/policiesProceduresStandardsStatements/rulesProcedures/13students/academic/13.99.99.R0.01.pdf

Academic Integrity

Students at Texas A&M University-Commerce are expected to maintain high standards of integrity and honesty in all of their scholastic work. For more details and the definition of academic dishonesty see the following procedures:

<u>Undergraduate Academic Dishonesty 13.99.99.R0.03</u> <u>Undergraduate Student Academic Dishonesty Form</u>

http://www.tamuc.edu/aboutUs/policiesProceduresStandardsStatements/rulesProcedures/documents/13.99.99.R0.03UndergraduateStudentAcademicDishonestyForm.pdf

<u>Graduate Student Academic Dishonesty Form</u>

http://www.tamuc.edu/academics/graduateschool/faculty/GraduateStudentAcademicDishonestyFormold.pdf

http://www.tamuc.edu/aboutUs/policiesProceduresStandardsStatements/rulesProcedures/13students/undergraduates/13.99.99.R0.03UndergraduateAcademicDishonesty.pdf

Students with Disabilities-- ADA Statement

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 Velma K. Waters Library Rm 162 Phone (903) 886-5150 or (903) 886-5835

Fax (903) 468-8148

Email: studentdisabilityservices@tamuc.edu

Website: Office of Student Disability Resources and Services

http://www.tamuc.edu/campusLife/campusServices/studentDisabilityResourcesAndServ

ices/

Nondiscrimination Notice

Texas A&M University-Commerce will comply in the classroom, and in online 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 <u>Carrying Concealed Handguns On Campus</u> 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 9-1-1.

A&M-Commerce Supports Students' Mental Health

The Counseling Center at A&M-Commerce, located in the Halladay Building, Room 203, offers counseling services, educational programming, and connection to community resources for students. Students have 24/7 access to the Counseling Center's crisis assessment services by calling 903-886-5145. For more information regarding Counseling Center events and confidential services, please visit www.tamuc.edu/counsel

COURSE OUTLINE / CALENDAR

Tentative Schedule:

- Part 1: Introduction, data input and output (Week 1-4)
- Part 2: Classification analysis 1 (Week 5)
- Midterm Exam (1): Week 6
- Part 3: Classification analysis 2 (Week 7-8)
- Part 4: Regression analysis (Week 9-10)
- Midterm Exam (2): Week 11
- Part 5: Clustering analysis (Week 12-13)
- Part 6: Advanced topics: class imbalance, outliers, ROC (Week 14)
- Course review: Week 15
- Final Exam: Week 16 (cumulative)