CS5270.Data Mining & Warehouses

COURSE SYLLABUS: FALL 2016

Instructor: Dr. Oner Ulvi Celepcikay
Time: n/a (Online Course)

Office Location: n/a (Online Course)
Office Hours: n/a (Online Course - contact me by email or cell phone)
Office Phone: (713) 966-9354 (cell phone)
University Email Address: celepcikayo@uhd.edu
Preferred Form of Communication: Email or Cell Phone for emergencies
Communication Response Time: 24 hours.

COURSE INFORMATION

Textbook Required

Required Book (DM): Introduction to Data Mining, by Pang-Ning Tan, Michael Steinbach and Vipin Kumar, 2006, 1st Edition, Addison-Wesley. Paper Book-ISBN 13: 978-0321321367

Software Required

I will use the Blackboard (https://bb.uhd.edu/) to provide you with online course materials including lecture notes, labs, and course announcements.

Course Description

The course presents comprehensive introduction to data mining techniques and provides the students with necessary background for the application of data mining to real problems. Topics include data processing, visualization, classification, predictive modeling, association analysis, clustering, and anomaly detection.

Student Learning Outcomes

After taking this course, a student should

1. Be able to understand strength and applicability of the data mining techniques. In particular, the classification techniques such as decision trees and clustering techniques including cover K-means clustering, agglomerative hierarchical clustering and DBSCAN.

2. Be able to understand data related issues including data types, data quality, and preprocessing steps to make the data suitable for data mining.

3. Be able to utilize basic visualization techniques and statistics for better understanding of the data at hand and the relationship among data attributes.

4. Be able to utilize the Python programming language for data analysis to develop data mining algorithms, and data preprocessing tools.

COURSE REQUIREMENTS

The course requires programming knowledge. Students are required to complete CS 3304 before taking this course. Students not meeting this prerequisite are subject to being dropped from the course without prior notification at their own expense.

We will use Python in this course but knowledge of Python is not a prerequisite. You need to have experience in at least one programming language; series of assignments will be given to assist you to learn Python and getting ready for team projects.

GRADING

Course Grade: Course grades will be determined as follows:

Activity	Weight
Exam-1 [DM Chapters 1, 2, 3, 4]	20%
Exam-2 [DM Chapter 4, 5, 8, 9]	20%
Final [All Covered Materials]	30%
Homework Assignments	30%

Your final course grade will be determined by following standard college formula:

A = 90%-100%

B = 80%-89%

C = 70%-79%

D = 60%-69%

F = 59% or Below

MAKE-UP POLICIES

Homework Assignments: Assignments are to be completed and turned in *by the due date*. **For each late day, 10 percent of the total possible points will be deducted.** No work will be accepted more than 5 days late.

Exams: Make-up exams will *only* be given in cases of documented emergencies. It is your responsibility to contact your instructor with documentation of your emergency.

Quizzes: No Make-ups for quizzes.

All missed grades will be recorded as zeros.

TECHNOLOGY REQUIREMENTS

- To fully participate in online courses you will need to use a current Flash enabled internet browser. For PC and Mac users the suggested browser is Mozilla Firefox.

- 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 courses 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.
  - *For courses utilizing video-conferencing tools and/or an online proctoring solution, a webcam and microphone are required.

- You will need to have Python 3x installed on your machine.
- Current anti-virus software must be installed and kept up to date.
- Run a browser check through the Pearson LearningStudio Technical Requirements website. Browser Check http://help.ecollege.com/LS_Tech_Req_WebHelp/en-us/#LS_Technical_Requirements.htm#Browset

- Running the browser check will ensure your internet browser is supported.
  - Pop-ups are allowed.
  - JavaScript is enabled.
  - Cookies are enabled.

- You will need some additional free software (plug-ins) for enhanced web browsing. Ensure that you download the free versions of the following software:
  - Adobe Reader https://get.adobe.com/reader/
  - Adobe Flash Player (version 17 or later) https://get.adobe.com/flashplayer/
  - Adobe Shockwave Player https://get.adobe.com/shockwave/
  - Apple Quick Time http://www.apple.com/quicktime/download/

- At a minimum, you must have Microsoft Office 2013, 2010, 2007 or Open Office. Microsoft Office is the standard office productivity software utilized by faculty, students, and staff. Microsoft Word is the standard word processing software, Microsoft Excel is the standard spreadsheet software, and Microsoft PowerPoint is the standard presentation software. Copying and pasting, along with attaching/uploading documents for assignment submission, will also be required. If you do not have Microsoft Office, you can check with the bookstore to see if they have any student copies.

• For additional information about system requirements, please see: System Requirements for LearningStudio https://secure.ecollege.com/tamuc/index.learn?action=technical

ACCESS AND NAVIGATION

Pearson LearningStudio (eCollege) Access and Log in Information
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
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.

Pearson LearningStudio Student Technical Support

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/

The student help desk may be reached in the following ways:

Chat Support: Click on *'Live Support'* on the tool bar within your course to chat with a Pearson LearningStudio Representative.

Phone: 1-866-656-5511 (Toll Free) to speak with Pearson LearningStudio Technical Support Representative

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.

Policy for Reporting Problems with Pearson LearningStudio

Should students encounter Pearson LearningStudio based problems while submitting assignments/discussions/comments/exams, the following procedure must be followed:

1. Students must report the problem to the help desk. You may reach the helpdesk at 1-866-656-5511.
2. Students must file their problem with the helpdesk and obtain a helpdesk ticket number
3. Once a helpdesk ticket number is in your possession, students should email me to advise me of the problem and provide me with the helpdesk ticket number.
4. I will call the helpdesk to confirm your problem and follow up with you

PLEASE NOTE: Your personal computer and internet access problems are not a legitimate excuses for filing a ticket with the Pearson LearningStudio Help Desk. Only Pearson LearningStudio based problems are legitimate reasons to contact the Help Desk. You strongly are encouraged to check for your internet browser compatibility **BEFORE** the course begins and take the Pearson LearningStudio tutorial offered for students who may require some extra assistance in navigating the Pearson LearningStudio platform.

myLeo Support
Your myLeo email address is required to send and receive all student correspondence. Please email 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

Learner Support
The One Stop Shop was created to serve you by providing as many resources as possible in one location. http://www.tamuc.edu/admissions/onestopshop/
The Academic Success Center provides academic resources to help you achieve academic success.
http://www.tamuc.edu/campusLife/campusServices/academicSuccessCenter/

FREE MobilE APPS

The Courses apps for phones have been adapted to support the tasks students can easily complete on a smaller device. Due to the smaller screen size course content is not presented. The Courses app is free of charge. The mobile Courses Apps are designed and adapted for different devices.

App Title:	**iPhone –** Pearson LearningStudio Courses for iPhone
	Android – LearningStudio Courses - Phone
Operating System:	**iPhone** - OS 6 and above
	Android – Jelly Bean, Kitkat, and Lollipop OS
iPhone App URL:	https://itunes.apple.com/us/app/pearson-learningstudio-courses/id977280011?mt=8
Android App URL:	https://play.google.com/store/apps/details?id=com.pearson.lsphone

Once downloaded, search for Texas A&M University-Commerce, and it should appear on the list. Then you will need to sign into the myLeo Mobile portal.

LearningStudio Notifications

Students can be alerted to course activities via text on their mobile phones or up to two email addresses.

Based on their preferences, students can automatically receive a push notification with every new: course announcement, threaded discussion post, grade, and/or assignment without having to login to the course. Enrolled students will automatically receive email notifications for announcements and can opt out of this feature. To receive text notifications, students must opt in.

To begin setting up notifications, go into your course in LearningStudio and click on the bell-shaped Notifications icon on the main menu ribbon.

By default the student's university email address will appear. This cannot be changed in LearningStudio. Additional email addresses may be added by clicking the Add button. After all of the other selections are completed be sure to click the Save and Finish button

COMMUNICATION AND SUPPORT

Interaction with Instructor Statement

The instructor's communication response time and feedback on assessments are stated clearly.

COURSE AND UNIVERSITY PROCEDURES/POLICIES

Student Conduct In Class Policy: Any acts of classroom disruption that go beyond the normal rights of students to question and discuss with instructors the educational process relative to subject content will not be tolerated, in accordance with the Academic Code of Conduct described in the Student Handbook.

Electronic Devices In Class Policy: Cellular phones, pagers, CD players, radios, and similar devices are prohibited in the classroom and laboratory facilities. Calculators and computers are prohibited during examinations and quizzes, unless specified. Reasonable laptop-size computers may be used in lecture for the purpose of taking notes.

University Specific Procedures
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.
http://www.tamuc.edu/admissions/registrar/documents/studentGuidebook.pdf
Students should also consult the Rules of Netiquette for more information regarding how to interact with students in an online forum: Netiquette
http://www.albion.com/netiquette/corerules.html

TAMUC Attendance

For more information about the attendance policy please visit the Attendance webpage and Procedure 13.99.99.R0.01.
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:

Undergraduate Academic Dishonesty 13.99.99.R0.03

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

Graduate Student Academic Dishonesty 13.99.99.R0.10

http://www.tamuc.edu/aboutUs/policiesProceduresStandardsStatements/rulesProcedures/13students/graduate/13.99.99.R0.10GraduateStudentAcademicDishonesty.pdf

ADA Statement
Students 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 or (903) 886-5835
Fax (903) 468-8148
Email: Rebecca.Tuerk@tamuc.edu
Website: Office of Student Disability Resources and Services
http://www.tamuc.edu/campusLife/campusServices/studentDisabilityResourcesAndServices

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 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 9-1-1.

COURSE OUTLINE / CALENDAR

#	Topic
1	**Course Introduction,** Syllabus Overview **What is Data Mining?** DM Chapter 1
2	**Data**, DM Chapter: 2
3	**Exploring Data DM Chapter 3**
4	**Classification & Decision Trees**, DM Chapter 4 - Measures of Node Impurity - Classification Issues
5	**- Exam I**
6	- Classification Model Evaluation - Comparing Classification Models
7	**Cluster Analysis,** DM Chapter 8 - Clustering Overview - K-means Clustering
8	- Hierarchical Clustering - DBSCAN Clustering - Cluster Evaluation
9	**Cluster Analysis II - DM Chapter 9**
10	**- Exam 2**
11	- Exam 2 Review - R: An introduction to popular statistical language*
12	- Regression Analysis (Subject to Change) **- Association Analysis** DM Ch 6 (Subject to Change)
13	- Team Presentations - Final Exam Review
14	**Final Exam**

*These items are outside materials and subject to change depending on how much of the material we cover.