CSCE 5073, Data Mining (3 credit hours), Elective

Catalog Description: Topics include data preprocessing; data warehousing and online analytical processing; data cube; mining frequent patterns, associations and correlations; supervised learning including decision tree induction, naïve Bayesian classification, support vector machine and K-nearest neighbor learning; unsupervised learning including K-means clustering and hierarchical clustering; outlier analysis; and data mining in cloud computing, social media, bioinformatics and healthcare applications.

Prerequisites: CSCE 3193 Programming Paradigms and (either INEG 2314 or STAT 3103) or CSCE graduate standing

Textbook/required material:

Data Mining: Concepts and Techniques, by Jiawei Han, Micheline Kamber, and Jian Pei, 3rd edition, 2011, Morgan Kaufmann Publishers. ISBN: 978-0123814791 (**required**)

Introduction to Data Mining, by Pang-Ning Tan, Michael Steinbach, and Vipin Kumar, 2006, Addison-Wesley, ISBN: 9780321321367 (**optional**)

Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman. Mining of Massive Datasets., 2014 (**optional**)

This course incorporates lectures and discussions of assigned readings.

Goals: The goal of the class is for students to 1) learn the core data mining technologies including data processing, frequent mining, classification and clustering, and 2) understand the advanced data mining techniques to conduct practical data analysis in various applications.

Topics covered:

- Introduction to data mining
- Data processing
- Data warehousing and online analytical processing
- Data cube technology
- Mining frequent patterns, associations and correlations
- Advanced frequent pattern mining
- Classification: basic concepts
- Classification: advanced methods
- Cluster analysis: basic concepts and methods
- Cluster analysis: advanced methods
- Outlier analysis and fraud detection
- Data mining in various applications

Class/laboratory schedule: Meets either 3 times a week for 50 minutes or 2 times a week for 1 hour 20 minutes for 15 weeks.

Relationship of course to Computer Engineering Program Student Outcomes:

- (a) An ability to apply knowledge of mathematics, science, and engineering.
- (b) An ability to design and conduct experiments, as well as to analyze and interpret data.

- (c) An ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability.
- (e) An ability to identify, formulate, and solve engineering problems.
- (k) An ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.

Relationship of course to Computer Science Program Student Outcomes:

- (a) An ability to apply knowledge of computing and mathematics appropriate to the discipline.
- (b) An ability to analyze a problem, and identify and define the computing requirements appropriate to its solution.
- (c) An ability to design, implement and evaluate a computer-based system, process, component or program to meet desired needs.
- (i) An ability to use current techniques, skills, and tools necessary for computing practices.
- (j) An ability to apply mathematical foundations, algorithmic principles, and computer science theory in the modeling and design of computer-based systems in a way that demonstrates comprehension of the tradeoffs involved in design choices.

CSCE 5073 – Data Mining Spring 2024

General Information

- Class website: http://csce.uark.edu/~xintaowu/5073/5073.htm
- Time/Location: Tuesday & Thursday, 3:30pm -4:45am, JBHT 239
- Instructor: Xintao Wu, JBHT 516, Email: xintaowu@uark.edu

Grading

Homework and in-class quiz 10%, Project 30%, Midterm 20%, Final 40%

Schedule

- http://csce.uark.edu/~xintaowu/5073/schedule.htm

Project description

- http://csce.uark.edu/~xintaowu/5073/proj.htm

Course Mode of Delivery

The course delivery mode will be face-to-face. While some asynchronous delivery may be involved, the bulk of content is delivered face-to-face during the regularly scheduled class time.

Attendance Policy

The University of Arkansas will primarily offer in-person instruction in the 2023-2024 academic year. Most of the university's academic programs have essential in-person components. Class attendance is the responsibility of each student and expected. If you are absent, it is your responsibility to obtain assignments, notes, and any class information given. If you must quarantine, self-isolate, or miss class during the semester because of COVID-19 or other illness, please contact the instructor via email and do not come to class. Class information will be provided to students who must miss class due to COVID-19 or other excused absences on a short-term basis. Contact the Center for Education Access (CEA) to determine if you think that you have a disability that permits you from participating in person.

Office Hours

2:30-3:30pm Thursday.

Academic Dishonesty Policy

As a core part of its mission, the University of Arkansas provides students with the opportunity to further their educational goals through programs of study and research in an environment that promotes freedom of inquiry and academic responsibility. Accomplishing this mission is only possible when intellectual honesty and individual integrity prevail. Each University of Arkansas student is required to be familiar with and abide by the University's 'Academic Integrity Policy' at honesty.uark.edu/policy. Students with questions about how these policies apply to a particular course or assignment should immediately contact their instructor.

Communication Plan

In this class our official mode of communication is through uark.edu email. Students are responsible for checking their UARK accounts regularly. All communication between student and instructor and between student and student should be respectful and professional.

Continuity Plan

In the event of an extended campus closure, we will use our Blackboard course as the portal for the delivery of course materials and office hour and UARK email for communications. Please check both of these areas immediately for guidance and directions from me.

Technology/Software Requirements

Access to a reliable Internet connection is required for this course. A problem with your Internet access may not be used as an excuse for late, missing, or incomplete coursework. If you experience problems with your Internet connection while working on this course, it is your responsibility to find an alternative Internet access point, such as a public library or Wi-Fi hotspot.

Unauthorized Use of Class Recordings

Instructors may record the class and make the class available to students through Blackboard. These recordings may be used by students ONLY for the purposes of the class. Students may not download, store, copy, alter, post, share, or distribute in any manner all or any portion of the class recording, e.g. a 5-second clip of a class recording sent as a private message to one person is a violation of this provision. This provision may protect the following interests (as well as other interests not listed): faculty and university copyright; FERPA rights; and other privacy interests protected under state and/or federal law. Failure to comply with this provision will result in a referral to the Office of Student Standards and Conduct for potential charges under the Code of Student Life. In situations where the recordings are used to gain an academic advantage, it may also be considered a violation of the University of Arkansas' academic integrity policy.

Unauthorized Recording by Student

Recording, or transmission of a recording, of all or any portion of a class is prohibited unless the recording is necessary for educational accommodation as expressly authorized and documented through the Center for Educational Access with proper advance notice to the instructor. Unauthorized recordings may violate federal law, state law, and university policies. Student made recordings are subject to the same restrictions as instructor-made recordings. Failure to comply with this provision will result in a referral to the Office of Student Standards and Conduct for potential charges under the Code of Student Life. In situations where the recordings are used to gain an academic advantage, it may also be considered a violation of the University of Arkansas' academic integrity policy.

Recording of Class Lectures

By attending this class, student understands the course may be recorded and consents to being recorded for official university educational purposes. Be aware that incidental recording may also occur before and after official class times.

Unauthorized Use and Distribution of Class Notes

Third parties may attempt to connect with you to buy your notes and other course information from this class. I will consider distributing course materials to a third party without my authorization a violation of my intellectual property rights and/or copyright law as well as a violation of the University of Arkansas' academic integrity policy. Continued enrollment in this class signifies your intent to abide by the policy. Any violation will be reported to the Office of Academic Initiatives and Integrity.

Please be aware that such class materials that may have already been given to such third parties may contain errors, which could affect your performance or grade. Recommendations for success in this course include coming to class on a routine basis, visiting me during my office hours, connecting with the Teaching Assistant (TA), and making use of Student Success Center. If a third party should contact you regarding such an offer, I would appreciate your bringing this to my attention. We all play a part in creating a course climate of integrity.

Changes to the Syllabus

The standards and requirements set forth in this syllabus may be modified at any time by the course instructor. Notice of such changes will be by announcement in class or by changes to this syllabus posted in Blackboard.

Emergency Procedures

Many types of emergencies can occur on campus; instructions for specific emergencies such as severe weather, active shooter, or fire can be found at emergency.uark.edu.

Severe Weather (Tornado Warning):

- Follow the directions of the instructor or emergency personnel
- Seek shelter in the basement or interior room or hallway on the lowest floor, putting as many walls as possible between you and the outside
- If you are in a multi-story building, and you cannot get to the lowest floor, pick a hallway in the center of the building
- Stay in the center of the room, away from exterior walls, windows, and doors

Violence / Active Shooter (CADD):

- CALL - 9-1-1
- AVOID - If possible, self-evacuate to a safe area outside the building. Follow directions of police officers.
- DENY - Barricade the door with desk, chairs, bookcases or any items. Move to a place inside the room where you are not visible. Turn off the lights and remain quiet. Remain there until told by police it’s safe.
- DEFEND - Use chairs, desks, cell phones or whatever is immediately available to distract and/or defend yourself and others from attack.

Inclement Weather Policy

- If the UofA is closed, then there will not be class. See http://www.uark.edu to determine if the UofA is closed.
- If the Fayetteville Public Schools are closed *because of the weather*, then there will not be class.
- If there is bad weather and the U of A and the Fayetteville Public Schools are open, I will post an announcement on the class website, which is typically Blackboard, or send by email.
- If you feel that you cannot safely come to class, then do not come. Use your own judgment.