

**CS 5593 DATA MINING  
FALL 2020**

**SECTION 001: 2:00 PM - 3:15 PM (Monday, Wednesday) via Zoom Video Conferencing  
(The Zoom meeting credentials for the lectures and the office hours are posted  
on the class website on Canvas)**

**Instructor:**

Dr. Le Gruenwald; Email: [ggruenwald@ou.edu](mailto:ggruenwald@ou.edu); Office Hours: 6:00 PM-5:00 PM Monday, Wednesday via Zoom

**Teaching Assistant:**

Siva Rama Krishan Ganta; Email: [shivaskr1234@ou.edu](mailto:shivaskr1234@ou.edu); Office Hours via Zoom are posted on the class website on Canvas

**Required Textbook:**

Pang-Ning Tang, Michael Steinbach, Anuij Karpatne, and Vipin Kumar, *Introduction to Data Mining, 2<sup>nd</sup> Edition*, Pearson Education, Inc., 2019

**Reference Materials:**

1. CRAN, *An Introduction to R*, <https://cran.r-project.org/doc/manuals/r-release/R-intro.html>
2. Ian H. Witten, Eibe Frank, and Mark A. Hall, "Data Mining – Practical Machine Learning Tools and Techniques," 3rd Edition, Chapters 10-17, Morgan Kaufmann, 2011 (*available online on OU Library Website*)
3. Jiawei Han and Micheline Kamber. *Data Mining: Concepts and Techniques, Third Edition*, Morgan Kaufman Publishers, 2012
4. Jure Leskovec, Anand Rajaraman and Jeffrey D. Ullmann, *Mining of Massive Datasets*, 2014
5. Additional materials provided by the instructor

**Prerequisite:**

Permission of instructor

**Course Website:**

<https://canvas.ou.edu>

**Course Coverage:**

This course covers the process, concepts and techniques in data mining, including advanced techniques that deal with big data. It provides students with the necessary background for the application of data mining to real world problems. The topics to be covered include:

| <b>Topic</b>   | <b>Reading</b>              |
|--|-----------------------------|
| 1. Introduction to Data Mining                                   | Ch. 1 (Text)                |
| 2. Data  | Ch. 2 (Text)                |
| 3. Data Mining Tools: WEKA and R                                 | Ref. 1, 2                   |
| 4. Exploring Data  | Ch. 3 (1 <sup>st</sup> Ed.) |
| 5. Classification: Concepts, Decision Trees and Model Evaluation | Ch. 3 (Text)                |
| 6. Association Analysis  | Ch. 5 (Text)                |
| 7. Cluster Analysis  | Ch. 7 (Text)                |
| 8. Anomaly Detection   | Ch. 9 (Text)                |
| 9. Classification: Alternative Techniques                        | Ch. 4 (Text)                |
| 10. Association Analysis: Advanced Concepts                      | Ch. 6 (Text)                |
| 11. Big Data Mining  | Ch. 8 (Text), Ref. 4        |
| 12. Avoiding False Discoveries                                   | Ch. 10 (Text)               |

**Grading:**

|   |     |
|---|-----|
| Group Project (Proposal, Progress Report, Final Report, and Presentation) | 50% |
| Homework  | 50% |
| Bonus Credit Poll Questions   | 10% |

**BONUS CREDIT POLL QUESTIONS:** Starting from the class of Monday, August 31, 2020, we will set up poll questions on Zoom at appropriate time points DURING the lectures for you to take to get BONUS CREDIT. The poll questions given in a particular class are based on the materials discussed in that class only. The total number of bonus points of all the questions given in all the remaining classes together from the Monday class of August 31, 2020 to the end of the semester is 100 points, which is 10% of the total grade. All questions will be scored equally. If you answer a question correctly, you will get a full score for that question. If you answer a question wrongly, you will get a 1/3 score for that question. If you do not answer a question, you will get a zero score for that question regardless of the reasons.

**COPYRIGHT POLICY:** Sessions of this course may be recorded or live-streamed by the individual faculty member. These recordings are the intellectual property of the individual faculty member and may not be shared or reproduced without the explicit, written consent of the faculty member. In addition, privacy rights of others such as students, guest lecturers, and providers of copyrighted material displayed in the recording may be of concern. Students may not share any course recordings with individuals not enrolled in the class, or upload them to any other online environment.

**ACADEMIC INTEGRITY:** Homework is individual work; they must be done by you only – no collaboration with anyone else is allowed. Group Project and Presentation are group work; they must be done by you and your assigned group members only – no collaboration with other groups or anyone else is allowed. Plagiarism will result in action as specified in the Academic Integrity Code at OU: [http://integrity.ou.edu/files/Academic\\_Misconduct\\_Code.pdf](http://integrity.ou.edu/files/Academic_Misconduct_Code.pdf). Consult also the following web page for a Student's Guide to Academic Integrity at OU: [http://integrity.ou.edu/students\\_guide.html](http://integrity.ou.edu/students_guide.html).

**REASONABLE ACCOMMODATION POLICY:** Students requiring academic accommodation should contact the Disability Resource Center for assistance at (405) 325-3852 or TDD: (405) 325-4173. For more information please see the Disability Resource Center website <http://www.ou.edu/drc/home.html>. Any student in this course who has a disability that may prevent him or her from fully demonstrating his or her abilities should contact me personally as soon as possible so we can discuss accommodations necessary to ensure full participation and facilitate your educational opportunities.

**RELIGIOUS OBSERVANCE:** It is the policy of the University to excuse the absences of students that result from religious observances and to reschedule examinations and additional required classwork that may fall on religious holidays without penalty.

**ADJUSTMENTS FOR PREGNANCY/CHILDBIRTH RELATED ISSUES:** Should you need modifications or adjustments to your course requirements because of documented pregnancy-related or childbirth-related issues, please contact me as soon as possible to discuss. Generally, modifications will be made where medically necessary and similar in scope to accommodations based on temporary disability. Please see [www.ou.edu/content/eoo/faqs/pregnancy-faqs.html](http://www.ou.edu/content/eoo/faqs/pregnancy-faqs.html) for commonly asked questions.

**TITLE IX RESOURCES:** For any concerns regarding gender-based discrimination, sexual harassment, sexual misconduct, stalking, or intimate partner violence, the University offers a variety of resources, including advocates on-call 24.7, counseling services, mutual no contact orders, scheduling adjustments and disciplinary sanctions against the perpetrator. Please contact the Sexual Misconduct Office 405-325-2215 (8-5, M-F) or OU Advocates 405-615-0013 (24.7) to learn more or to report an incident.