

# CSCE 4380 / 5380 - Data Mining

## Instructor Information

Dr. Ryan Garlick, [garlick@unt.edu](mailto:garlick@unt.edu)

Office Hours: via Zoom link (times listed in Canvas). Please email at least 24 hours prior and I will send a Zoom meet link.

Teaching Assistants:

Dinesh Ravula [dineshreddyravula@my.unt.edu](mailto:dineshreddyravula@my.unt.edu)

Rana Ghoneim [ranaghoneim@my.unt.edu](mailto:ranaghoneim@my.unt.edu)

Office Hours: via Zoom link (times listed in Canvas). Please email at least 24 hours prior.

## Course Description, Structure, and Objectives

This course focuses on fundamental concepts, principles and techniques related to data mining. Data Mining refers to the process of automatic discovery of interesting patterns and knowledge from large data collections, including databases, data warehouses, document collections, and data streams. We will study important topics of data mining, including data preprocessing, frequent pattern and association rule mining, correlation analysis, classification and prediction, and clustering. Students will become acquainted with both the strengths and limitations of various data mining techniques.

Prerequisite: CSCE 3110.

This course is online and will follow a weekly module format.

Upon successful completion of this course, students will be able to:

1. Gain familiarity with key data visualization and data pre-processing methods
2. Be able to confidently apply both supervised and unsupervised methods across a wide range of real-world scenarios
3. Gain a fundamental understanding of time series prediction and its applications
4. Understand the basic principles of spatial data mining and its applications
5. Gain in depth experience in applying a major language used in data mining

## Required/Recommended Materials

All materials for the course are online and presented through Canvas. There is no textbook for the course.

## ADA Notice

The University of North Texas makes reasonable academic accommodation for students with disabilities. Students seeking reasonable accommodation must first register with the Office of Disability Access (ODA) to verify their eligibility. If a disability is verified, the ODA will provide you with a reasonable accommodation letter to be delivered to faculty to begin a private discussion regarding your specific needs in a course. You may request reasonable accommodations at any time; however, ODA notices of reasonable accommodation should be provided as early as possible in the semester to avoid any delay in implementation. Note that students must

obtain a new letter of reasonable accommodation for every semester and must meet with each faculty member prior to implementation in each class. Students are strongly encouraged to deliver letters of reasonable accommodation during faculty office hours or by appointment. Faculty members have the authority to ask students to discuss such letters during their designated office hours to protect the privacy of the student. For additional information, refer to the [Office of Disability Access](http://www.unt.edu/oda) website (<http://www.unt.edu/oda>). You may also contact ODA by phone at (940) 565-4323.

## Assessing Your Work

Homework	35%
Quizzes	35%
Midterm	15%
Final Exam	15%

A = 90-100%

B = 80-89%

C = 70-79%

D = 60-69%

F = 59% or below

Late work is not accepted and there is no extra credit or dropped assignments or quizzes.

## Course Schedule (subject to change)

Week 1: Introduction and Tools  
Week 2: Data Mining Fundamentals 1  
Week 3: Data Mining Fundamentals 2  
Week 4: Data Mining Fundamentals 3  
Week 5: Python Review for Data Science  
Week 6: Probability for Data Science  
Week 7: Data Pre-Processing  
Week 8: Midterm  
Week 9: Spring Break  
Week 10: Time Series  
Week 11: Spatial Data Mining  
Week 12: Association Analysis  
Week 13: Clustering / Classification  
Week 14: Stylometry  
Week 15: Project  
Week 16: Final Exam

Graduate students will have additional project requirements.

## Attendance and Participation

You are expected to keep pace with the class and turn assignments in when due (typically weekly). Materials will be posted well before the due date and you are encouraged to work ahead in case unforeseen circumstances arise.