University of North Texas
Department of Computer Science and Engineering

CSCE 4380/5380 – Data Mining

Session: Spring/2026

Classes:
4380-001 (6861) - MoWe 1:00PM - 2:20PM - NTDP E266
5380-001 (6111) - MoWe 1:00PM - 2:20PM - NTDP E266
5380-600 (14255**) -** MoWe 1:00PM - 2:20PM - NTDP E266
5380-002 (7324) - TuTh 2:30PM - 3:50PM - NTDP K120

Instructor:
Dr. Moawia Eldow
Office: E250E
Email: **moawia.eldow@unt.edu**
Office hours: Tu/Th 12:30-2:30 PM

IAs/TAs:

4380-001/5380-001,600 (MoWe):
Name: Sai Mahesh Mudavat
Email: SaiMaheshMudavat@my.unt.edu
Office & Office hours: TBA

5380-002 (TuTh):
Name: Md Fahimul Kabir Chowdhury
Email: MdFahimulKabirChowdhury@my.unt.edu
Office & Office hours: TBA

Name: Aniv Chakravarty
Email: AnivChakravarty@my.unt.edu
Office & Office hours: TBA

Description:
This course focuses on fundamental concepts, principles and techniques related to data mining. We will study important topics of data mining, including data preprocessing, frequent pattern and association rule mining, classification, clustering, anomaly detection, and some recent developments and trends.

Prerequisites:
Programming with one of the high-level languages such as C, C++, or Java; Introductory courses on data structures and algorithms, linear algebra and probability theory.

Learning Outcomes:
By the end of the course, students will

1. Be familiar with basic concepts and main tasks of data mining.
2. Be familiar with data types and data preprocessing.
3. Be able to apply the basic classification analysis including decision trees, rule-based classifiers, Bayesian classifiers, k-nearest neighbor, logistic regression, support vector machines, and artificial neural networks.
4. Be familiar with model overfitting, selection & evaluation and class imbalance problems.
5. Be able to apply the basic cluster analysis.
6. Be able to apply the basic association analysis.
7. Be able to apply the anomaly detection.
8. Be familiar with recent trends on data mining.

Textbooks:

1- Introduction to Data Mining, 2nd Edition, by Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. (IDM – Chapters 1, 2, 3, 4, 5, 7 and 9)
2- Data Mining: Concepts and Techniques, 4th Edition, by Jiawei Han, Micheline Kamber and Jian Pe. (DMCT – Chapter 12)

Grades and grading policy:

Participation	**5%**
Quizzes	**10%**
Homework Assignments	**30%**
Project	**15% (*5380 – with presentation, 4380 – without presentation*)**
Midterm Exam	**20%**
Final Exam	**20%**

The letter grade will be assigned based on the following scale:

Grade	4380/5380
A	90 and Above
B	[80-90)
C	[70-80)
D	[60-70)
F	Below 60

Class Participation:

Student participation in this class can be counted from the attendance, the Canvas contribution (the volume of student work/activity on Canvas) and the discussion participation in the class.

Quizzes:

Generally, quizzes will be given based on the readings and topics covered in classes. Quizzes will test students' knowledge on the most important aspects of the readings and understanding of the topics covered in some modules. ***Quizzes will be available only on Fridays.***

Homework Assignments:

Written individual homework assignments/exercises ***will be due at 11:59 p.m. on Fridays.*** Assignments must be turned in using the dropbox on canvas.

Project:

Group project must be submitted and may be presented in the class (3 to 5 students per project). Presentation of the projects is ***mandatory*** for graduates (**CSCE5380**) and ***optional*** for undergraduates (**CSCE4380**), and ***they will grant a bonus when they complete it***. Components of the project include a proposal and final project report. Projects must be turned in using the dropbox on canvas.

Exams:

There will be one midterm exam during the semester at the normal lecture time, which will cover the first half of the class topics. There will also be a final exam during finals week, which will cover the second half of the class topics.

Late Submission Policy:

Assignments may be turned in late, but not more than two weeks. All the late submissions **may lose a percentage of their graded point values** according to the following schedule:

On time: **0%,** 1-3 days: **10%,** 4-7 days: **20%,** 8-14 days: **40%,** > 14 days: **100%**

CSCE 4380/5380 - Course Outline (*Tentative Schedule*):

Week	**Reading chapters and Topics**	**Quizzes (Q), Homework (HW) (On/Due dates)**
Jan 12-16	Overview of Class Ch1 – Introduction	
Jan 19-23	**01/19** - *MLK Holiday* ***(No classes)*** **(*for 4380-1/5380-1, 600 only*)** Ch2 – Data & Data Preprocessing	**HW1 (01/30)** **Project (*Instructions posted on Canvas*)**
Jan 26-30	Ch3 – Basic Concepts & Decision Trees	**HW2 (02/06)**
Feb 2-6	Ch3 – Model Overfitting & Evaluation	***Q1 (02/06)***
Feb 09-13	Ch4 – Rule-based Classifiers	**HW3 (02/20)** **Project (*Proposal & groups are due on 02/13*)**
Feb 16-20	Ch4 – Naïve Bayes Classifier & Belief Networks	**HW4 (02/27)**
Feb 23-27	Ch4 – K-Nearest Neighbor & Logistic Regression	***Q2 (02/27)***
Mar 2-6	**Mid-Term Exam:** ***03/02/2026*** *(*MoWe class - 4380-1/5380-1,600*), **Available from 1:00 pm to 6:00 pm*** ***03/03/2026*** (TuTh class – 5380-2), ***Available from 2:00 pm to 7:00 pm***	
Mar 09-13	Spring Break **(*No classes*)**	
Mar 16-20	Ch4 - Artificial Neural Networks	
Mar 23-27	Ch4 – Support Vector Machine & Class Imbalance Problem	***Q3 (03/27)*****, HW5 (04/03)**
Mar 30-Apr 3	Ch7 – Basic Cluster Analysis	**HW6 (04/10)**
Apr 6-10	Ch5 – Basic Association Analysis	**HW7 (04/17)**
Apr 13-17	Ch9 – Anomaly Detection	***Q4 (04/17),*** **Project (*Final reports are due on 04/17*)**
Apr 20-24	Ch12 (DMCT) - Trends on Data Mining **Project Presentation Forum** ***04/22/2026 (***MoWe class - 4380-1/5380-1, 600***), Available from 1:00 pm to 11:59 pm*** ***04/23/2026 (***TuTh class – 5380-2***), Available from 1:00 pm to 11:59 pm***	
Apr 27-May 1	Review Week	
May 4-8	**Final Exam:** ***05/04/2026 (***MoWe class - 4380-1/5380-1, 600***), Available from 1:00 pm to 6:00 pm*** ***05/05/2026 (***TuTh class – 5380-2***), Available from 1:00 pm to 6:00 pm***	

Communication

Students are expected to communicate with the instructor/TAs through the office hours, using only the UNT emails of students and instructor/TAs, and/or the communication facilities provided in Canvas. Connect with me through email and/or by attending office hours. During busy times, my inbox becomes rather full, so if you contact me and do not receive a response within two business days, please send a follow up email. A gentle nudge is always appreciated.

Laboratory Safety Procedures and Guidelines policy

Students can access this policy at: Laboratory Safety Procedures and Guidelines policy (PDF) (https://policy.unt.edu/sites/default/files/06.049_Standard%20Syllabus%20Policy%20Statements_supplement.pdf).

Attendance and Participation

Research has shown that students who attend class are more likely to be successful. You should attend every class unless you have a university excused absence such as active military service, a religious holy day, or an official university function as stated in the Student Attendance and Authorized Absences Policy (PDF) (https://policy.unt.edu/sites/default/files/06.039_StudAttnandAuthAbsence.Pub2_.19.pdf). If you cannot attend a class due to an emergency, please let me know. Your safety and well-being are important to me.

Digital Requirements

This course has digital components. To fully participate in this class, students will need internet access to reference content on the Canvas Learning Management System. If circumstances change, you will be informed of other technical needs to access course content.
Information on how to be successful in a digital learning environment can be found at Learn Anywhere (https://online.unt.edu/learn).

How to Succeed in this Course

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 website (http://www.unt.edu/oda). You may also contact ODA by phone at (940) 565-4323.

Supporting Your Success and Creating an Inclusive Learning Environment

Every student in this class should have the right to learn and engage within an environment of respect and courtesy from others. We will discuss our classroom's habits of engagement and I also encourage you to review UNT's student code of conduct so that we can all start with the same baseline civility understanding (Code of Student Conduct) (https://deanofstudents.unt.edu/conduct)

Academic integrity, student misconduct, and professionalism policies for CSE

These are the policies of the Department of Computer Science and Engineering (CSE) at University of North Texas related to academic integrity, student misconduct, and professionalism.

All department policies on Academic Integrity and Student Conduct apply for this course – these are available at the following link:

https://engineering.unt.edu/cse/students/resources/academic-integrity.html

You can also download one copy form the course site on Canvas.

Artificial Intelligence in Academic Integrity

Students need to be aware the "unauthorized" use of any person or technology that assists in a student's assignment, project, or paper is considered cheating under the UNT Student Academic Integrity Policy (UNT Policy 6.003). Unless a professor or instructor gives explicit "authorization," AI cannot be used to assist in the completion of assignments, projects, or papers. Doing so will result in a "cheating" violation. Again, if uncertain contact the instructor prior to using AI tools.

Other University Polices

Students will be aware about Eagle Alert if there is a campus closing that will impact a class and the calendar is subject to change, refer to the Emergency Notifications and Procedures Policy (PDF) and the Campus Closures Policy (https://policy.unt.edu/policy/15-006).

Students can access these policies in Navigate (Navigate.unt.edu), in Canvas under the Help menu, in EIS, and on the Student Support Services & Policies page.