DSCI 4520 – Introduction to Data Mining
Spring 2023

“What I cannot create, I do not understand.”
Richard Feynman

Course Information

Course: DSCI 4520.002
Course Name: Introduction to Data Mining
Class Location: BLB 170
Time: Mo 2:00 pm – 4:50 pm
Website: http://learn.unt.edu

Instructor Contact Information

Instructor: Dr. Mahdi Ahmadi
Office: BLB 331C
Phone: 940-565-2946
Email: Mahdi.Ahmadi@unt.edu
Office Hours: Mo 12:30 pm – 2:00 pm
Tu 10:00 am – 12:30 pm
or by appointment
Zoom link: https://unt.zoom.us/j/3579223579

Course Description

How can data be utilized to make better decisions and solve real-life problems? How can data-driven information shape and support business decisions? What is the true value of data for businesses? How to turn data into solutions for business problems? This course is designed to address these and similar questions by focusing on extracting relevant and useful information from large databases and designing data-driven decision support systems through data mining techniques. In this course, we introduce technologies, processes, and practices to explore and investigate business data to gain insights for planning. The course will start with concepts and techniques for data preparation, summarization, visualization, and statistical analysis. Then, we will introduce methods such as decision trees, neural networks, association rule extraction, and classification ensembles to discover relationships and patterns in data that shed light on business problems. This course will examine principles and techniques of predictive analytics that can be used to shape the present and future of businesses.
Learning Objectives

1- To learn basic concepts and techniques of data mining such as exploratory analysis, data pre-processing, linear regression models, classification models, clustering, model performance evaluation, and neural networks.
2- To understand business problems, transform them into data mining problems, solve the problems, interpret results in the business context, and present the solutions effectively.
3- To identify when and how each data mining technique can be used to solve business problems.
4- To acquire a practical understanding of preparing and evaluating predictive models and explorative algorithms

Textbook, Materials, and Resources

Textbooks

  - Note: On the DMBA textbook website, data sets are available for assignment and practice: https://www.dataminingbook.com/book/r-edition

  - Note: An electronic version of this book is available to access through the UNT Library resources: https://discover.library.unt.edu/catalog/b6041002

I may also post educational videos, scholarly papers, blog posts, articles, and other materials on Canvas. Students should check their Canvas account regularly for additional materials.

Software

This is a hands-on course and students will learn data mining techniques through practicing them. We will use two main software in this course: Analytic Solver (previously known as XLMiner) and RapidMiner. Both software packages are freely available for the students in this course. Detailed instructions on how to access, install, and activate Analytic Solver and RapidMiner will be available on Canvas.

Delivery Method and Technology Requirements

The course will be delivered in person in the BLB 155 classroom. Active class participation is crucial to the learning and success of the students. Active participation means participation in the class discussions, presenting your team project in class, and engaging in constructive dialogues. I will deliver a relatively short lecture in each class, and the rest of the class will be discussions, students’ presentations, case studies, and hands-on practices. We will rely heavily on Canvas for the delivery of materials, practices, assignments, exams, and projects.
We will rely heavily on Canvas for the roster call, delivery of materials, assignments, exams, and the term project.

A laptop is required during all classes and exams. We will use a variety of analytical and data tools in the class. It is the student’s responsibility to have a working and functional laptop during class time and exams. Laptops can be checked out from the UNT Library. You can remotely connect to the UNT MyLab virtual computer lab and use the tools. To access UNT MyLab go to [https://mylab.unt.edu/](https://mylab.unt.edu/)

During office hours I will be available both in my office and on Zoom. Students can ask for an appointment outside of office hours. All communications with me should ONLY be done through my official UNT email address or Canvas

**Attendance Policy**

Students are expected to attend class meetings regularly and to abide by the attendance policy established for the course. It is important that you communicate with the professor and the instructional team prior to being absent, so you, the professor, and the instructional team can discuss and mitigate the impact of the absence on your attainment of course learning goals. Please inform the professor and instructional team if you are unable to attend class meetings because you are ill, in mindfulness of the health and safety of everyone in our community.

---

**Course Schedule**

The following table shows the tentative schedule for the course. **Dates, topics, and activities are subject to change.**

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Subject</th>
<th>Reading</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16-Jan</td>
<td>MLK day. No Class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>23-Jan</td>
<td>Course syllabus + Introduction to Data Mining</td>
<td>Syllabus + Ch 1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>30-Jan</td>
<td>Review of Business Statistics in Excel &amp; Data</td>
<td>Chapter 2</td>
<td>Practice 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mining Process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6-Feb</td>
<td>Data Preparation, Visualization, and Exploratory</td>
<td>Chapter 3</td>
<td>Practice 2, Quiz 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Analytics I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13-Feb</td>
<td>Linear Regression</td>
<td>Chapter 6</td>
<td>Practice 3</td>
</tr>
<tr>
<td>6</td>
<td>20-Feb</td>
<td>Linear Regression cont. + k-Nearest Neighbors</td>
<td>Chapters 6 &amp; 7</td>
<td>Practice 4, Quiz 2</td>
</tr>
<tr>
<td>7</td>
<td>27-Feb</td>
<td>Exam 1 - Project first milestone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>6-Mar</td>
<td>Cluster Analysis and Segmentation</td>
<td>Chapter 15</td>
<td>Practice 5, Quiz 3</td>
</tr>
<tr>
<td>9</td>
<td>13-Mar</td>
<td><em>Spring Break -- No Class</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>20-Mar</td>
<td>Introduction to Classification &amp; Evaluating</td>
<td>Chapter 5</td>
<td>Practice 6, Quiz 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classifiers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>27-Mar</td>
<td>Naïve Bayes Classifier</td>
<td>Chapters 8</td>
<td>Practice 7</td>
</tr>
</tbody>
</table>
12 3-Apr Logistic Regression Chapter 10 Practice 8, Quiz 4
13 10-Apr Exam 2 - Project second milestone
14 17-Apr Decision Trees & Random Forests Chapter 9 Practice 9
15 24-Apr Association Rules & Recommender Systems Chapter 11 Practice 10
16 1-May Neural Networks Chapter 11 Quiz 5
16* 4-May Course Review and Project Presentations
17 8-May Final exam & final project report

* Because one class day is missing in the beginning of the semester (MLK Day) I have added an extra class meeting in Week 16, during which we will review the course and students will present their group projects. Those who cannot attend this extra meeting, due to a conflict, have the option to participate remotely. The exact date will be decided later, and the details will be posted on Canvas.

**Grade Components**
Grades will be based on student performance in the course, exams, assignments, term project, and class participation.

<table>
<thead>
<tr>
<th>Assessment Component</th>
<th>Points</th>
<th>Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice Assignments</td>
<td>300</td>
<td>30%</td>
</tr>
<tr>
<td>Pop Quizzes</td>
<td>100</td>
<td>10%</td>
</tr>
<tr>
<td>Exams</td>
<td>350</td>
<td>35%</td>
</tr>
<tr>
<td>Term Project</td>
<td>250</td>
<td>25%</td>
</tr>
</tbody>
</table>

**Practice Assignments**
This is a hands-on course in which you would learn how to create, use, and interpret data mining techniques. Throughout the course, 10 practice assignments are designed to help you learn the techniques and concepts. I will explain and initiate each assignment in the class and help students start working on them. After initiating each assignment in the class, students will have one week to complete and submit it.

**Pop Quizzes**
The purpose of pop quizzes is to assess and establish your learning in a short period. The tentative number and dates of quizzes are shown in the class schedule table and are subject to change. All quizzes will be open-book and should be taken individually in the class.

**Exams**
Three closed-book and closed-note exams will be given (two exams in the middle of the semester and a final exam). The final exam composes 15% of your final grade and will be a comprehensive evaluation of everything you’ve learned in this course. The dates of the exam are shown in the class schedule table. All exams should be taken in the classroom during the normal class meeting time through LockDown Browser. The final exam will be given on May 8 at 1:30 pm (following the university Final Exam Schedule). Students will be individually tested over all materials covered in the lectures, course readings, assignments, case studies, and tutorials.
Make-up exams will NOT be given in general. If you have a UNT-authorized valid and verified excuse for missing an exam, please communicate it with me at least 48 hours before the exam date. Requests for a make-up AFTER the exam date will only be accepted to review within 48 hours and it requires extraordinary circumstances with valid and verified documents. No request will be reviewed after 48 hours.

**Term Project**
This is an opportunity for students to work in a team to solve a business problem with the skills they learn in this course. Students can choose their topic and data. There will be two milestones (indicated in the class schedule table) to assess and review the progress of the project. Detailed instructions and rubrics will be posted on Canvas after the semester starts.

**Final Grades Policy**

<table>
<thead>
<tr>
<th>Final score</th>
<th>Final letter grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>90% &lt;</td>
<td>A</td>
</tr>
<tr>
<td>80% - 89%</td>
<td>B</td>
</tr>
<tr>
<td>70% - 79%</td>
<td>C</td>
</tr>
<tr>
<td>60% - 69%</td>
<td>D</td>
</tr>
<tr>
<td>&lt; 60%</td>
<td>F</td>
</tr>
</tbody>
</table>

Please read Instructor Policies and UNT Policies sections on academic integrity carefully.

**Grade Disputes**
Students have the right and are welcome to review assignments and exams after grading to improve their understanding of course material and check for the presence of grading errors. This review must be conducted within 7 days after the grades of the assignment or exam are posted on Canvas. All requests should be sent directly to my UNT email only with a clear and concise explanation of why and where you believe there has been a grading error. Grade appeals are to ensure mistakes do not negatively impact your grades. They are not intended to ensure you receive your desired final grade. Requests for blanket reconsideration of your graded deliverables will not be reviewed or answered. No changes to grades will be made after this period.

---

**Instructor Policies**

*Class Participation and Code of Conduct*
All students are expected to actively participate in all classes. Students are strongly encouraged to ask questions during my lectures and other students’ presentations.

Student behaviors that interfere with the instructor’s ability to conduct a class or other students’ opportunity to learn are unacceptable and disruptive and will not be tolerated in any instructional forum at UNT. Students engaging in unacceptable behavior will be directed to leave the classroom and the instructor may refer the student to the Dean of Students to consider whether the student’s conduct violated the Code of Student Conduct. The university’s expectations for student conduct apply to all
instructional forums, including university and electronic classrooms, labs, discussion groups, field trips, etc.

As a courtesy to the class, your classmates, and your instructor, you are asked to set your cell phones to vibrate. In the case of a personal emergency that requires you to answer your phone, you are asked to step out of the classroom.

**Academic Integrity**
I take academic integrity and honesty extremely seriously for all course activities such as assignments, exams, and the term project. This course is an excellent opportunity to practice what is expected from you as a business professional regarding integrity, trust, and honesty. All students are expected to maintain the highest standards of ethical and professional conduct when taking exams and doing assignments and the term project.

In this course, unless otherwise stated, individual work should be completed alone and using only resources explicitly outlined in the instructions. While external research may be permitted in assignment instructions, utilizing resources such as Chegg, Course Hero, and similar websites is not appropriate and is expressly forbidden. What may appear to be a relatively minor step outside the bounds of acceptable behavior can have a monumental impact on success within your academic program and beyond.

Students caught cheating or plagiarizing will receive a zero for that particular assignment or exam for the first time. Engaging in such behavior for the second time will result in an F for the course, with absolutely no exception. Additionally, the incident may be reported to the Dean of Students, who may impose further penalties or sanctions ranging from admonition to expulsion from the university.

Read the definition of cheating by the UNT carefully: [https://policy.unt.edu/policy/06-003](https://policy.unt.edu/policy/06-003)

**Communication**
Canvas is the primary channel for announcements, changes in due dates or the syllabus, new materials, grades, etc. It is the student’s responsibility to check their Canvas account for updates.

If a student wants to initiate communication for any purpose, they should do it with their UNT email to my UNT email: Mahdi.Ahmadi@unt.edu. Emails from non-UNT email systems will not be read. You can expect my response within 24 hours. Virtual appointments will be held on Zoom.

I value your feedback greatly and I strongly encourage you to not hesitate to communicate with me about any issues that you may have or see in the course. To make it easier, I will create an anonymous electronic drop box where you can leave your suggestions, opinions, and questions anonymously. Details will be posted on Canvas.

**Changes in the Syllabus**
The materials presented in this syllabus are tentative. I reserve the right to change and improve the materials and requirements as the semester unfolds, with sufficient announcements concerning exams, and assignments.
Late Work
I will not accept late work in this course. All work turned in after the deadline will receive a grade of zero unless the student has a university-authorized absence and provides documentation 48 hours before the deadline.

University Policies and Other Important Information

Academic Integrity Policy
The University of North Texas promotes the integrity of learning and embraces the core values of trust and honesty. Academic integrity is based on educational principles and procedures that protect the rights of all participants in the educational process and validate the legitimacy of degrees awarded by the University. In the investigation and resolution of allegations of student academic dishonesty, the University’s actions are intended to be corrective, educationally sound, fundamentally fair, and based on reliable evidence. Please read the full text of the UNT policy here.

Disability Accommodation Policy
The University of North Texas (UNT or University) does not discriminate on the basis of disability in admission, treatment, or access to its programs or activities, nor in employment in its programs or activities. The University is committed to providing equal educational access for qualified students with disabilities in accordance with state and federal laws, including the Americans with Disabilities Act of 1990 as Amended, and Section 504 of the Rehabilitation Act of 1973. In addition, the University is committed to making all programs and activities sponsored by UNT accessible, as required by the Texas Accessibility Standards and the Americans with Disabilities Act Accessibility Guidelines. To this end, all academic units are willing to make reasonable and appropriate adjustments to the classroom environment and the teaching, testing, or learning methodologies in order to facilitate equality of educational access for persons with disabilities.

Please read the full text of the UNT policy here. You can learn more about the UNT Office of Disability Access on their website or through their email (APPLY.ODA@UNT.EDU) or their phone number: 940-565-4323

Academic Deadlines
Dates of drop deadlines, final exams, etc., are published in the university calendar and the schedule of classes. Please make yourself aware of these dates.

Emergency Notification and Procedures
Eagle Alert is UNT’s official, campus-wide emergency notification system for emergency events, inclement winter weather closures, or Tornado Warnings. Eagle Alert allows UNT administrators to quickly contact campus community members by phone, text, and email. Eagle Alert will also post to the Eagle Alert Twitter. In addition to receiving direct messages by phone and email, the system includes a feature called desktop override that takes control over most UNT-owned computers in offices, classrooms, and public spaces on campus. During the test or in the event of an emergency, computer screens, presentation screens, and digital signs will display a full-screen alert.

UNT faculty, staff, and students are automatically enrolled in Eagle Alert. Individuals should remember, however, that if their personal contact information changes, they should go to my.unt.edu
(students/faculty) or my.untsystem.edu (staff) to update their information. Instructions for updating your information can be found here.

**Retention of Student Records**
All university records must be retained and disposed of in accordance with state law and the university’s record retention schedule as approved by the Texas State Library and Archives Commission. Please read the full text of the UNT policy [here](#).

**Student Success**
UNT endeavors to offer you a high-quality education and to provide a supportive environment to help you learn and grow. And, as a faculty member, I am committed to helping you be successful as a student. Here’s how to succeed at UNT: Show up. Find Support. Get advised. Be prepared. Get involved. Stay focused. To learn more about campus resources and information on how you can achieve success, go [http://success.unt.edu/](http://success.unt.edu/).

**Incomplete Grade (I)**
The grade of I (incomplete) is not given except for rare and very unusual emergencies, as per University guidelines. An I grade cannot be used to substitute for your poor performance in class. If you think you will not be able to complete the class satisfactorily, please drop the course.

**Campus Closures**
Should UNT close campus, it is your responsibility to check your official UNT e-mail account (EagleConnect) and Canvas to determine if your instructor plans to modify class activities, and how. This may include changing assignment due dates, rescheduling exams, holding online classes, etc.

**Student Evaluation of Instruction**
Student feedback is important and an essential part of participation in this course. The student evaluation of instruction is a requirement for all organized classes at UNT. The survey will be made available during weeks 13, 14 and 15 of the long semesters to provide students with an opportunity to evaluate how this course is taught. Students will receive an email from UNT SPOT Course Evaluations with the survey link. Students should look for the email in their UNT email inbox. Simply click on the link and complete the survey. For additional information, please visit the [SPOT website](http://success.unt.edu/).