Course Description
Introduction to data mining which includes main data mining tasks, e.g., classification, clustering, association rules, and outlier detection, and some of the latest developments, e.g., mining temporal and spatial data.

Learning Outcomes
By the end of the course you will

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

Course Requirements
Attendance: Physical attendance is mandatory. Lectures, videos, and class discussions will contain vital information needed to do well on exams.

There are two textbooks that are highly recommended for the R language. Both are available online free of charge.

1. Hashler, Michael; An R Companion for Introduction to Data Mining, available from Preface | An R Companion for Introduction to Data Mining (mhahsler.github.io)
The major reference for the lecture content will be the power point slides given on a weekly basis. Any additional online reading will be referenced online.

**Schedule**

<table>
<thead>
<tr>
<th>Week #</th>
<th>Lecture Topic</th>
<th>Tutorial</th>
<th>Assessment/Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>Introduction to Data Mining; Data Pre-processing</td>
<td>Tutorial 1 Data Pre-processing</td>
<td></td>
</tr>
<tr>
<td>Week 2</td>
<td>Rule Based and Decision Tree Classification</td>
<td>Tutorial 2: Rule Based and Decision Tree Classification</td>
<td>Project 1 handed out</td>
</tr>
<tr>
<td>Week 3</td>
<td>KNN and Neural Network Classifiers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 4</td>
<td>Association Rule Mining</td>
<td>Tutorial 3: Association Rule Mining</td>
<td></td>
</tr>
<tr>
<td>Week 5</td>
<td>Preparation for Midterm on Monday (<em>this will be in the form of a video, as July 4 is a Public Holiday</em>)</td>
<td>Midterm exam (online) on Wednesday in class</td>
<td>Project 1 due</td>
</tr>
<tr>
<td>Week 6</td>
<td>Clustering 1</td>
<td>Tutorial 4: Clustering</td>
<td>Project 2 handed out</td>
</tr>
<tr>
<td>Week 7</td>
<td>Clustering 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 8</td>
<td>Time Series Analysis</td>
<td>Tutorial 5: Time Series Analysis</td>
<td></td>
</tr>
<tr>
<td>Week 9</td>
<td>Spatial Data Mining</td>
<td></td>
<td>Project 2 due</td>
</tr>
<tr>
<td>Week 10</td>
<td>Preparation for Final Exam on Monday Final exam (online) on Wednesday 10 August in class</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Evaluation**

Midterm* and Final Exam, 25% each  
Projects (1, 2): 25% each

For the midterm test two optional quizzes will be offered in class in Week 7 (Monday and Wednesday) that will give you an opportunity to boost your score up to a maximum of 7 marks (i.e., a maximum boost of 10%).
Grade ranges A: 90-100; B: 80-89; C: 70-79; D: 60-69; F: <60 (raw final grades will be curved)

Prerequisites: Probability theory, Calculus, Data structures, Proficiency with C/C++, Java, Python, or Matlab. Note that this course will use R in the programming assignments.

**Course Policies and Procedures**

1. Emails will be answered as promptly as possible. Emails outside normal working hours (8 am to 5 pm) will be answered on the next working day.

2. Work handed in for grade (homework, project report, etc.) **MUST BE YOUR OWN effort only**. Students are **NOT** allowed to use online solutions from previous course offerings, websites, etc. **This will be strictly checked and enforced.** The students should adhere to the UNT policies and procedures on the Code of Academic Integrity. 06.003 Student Academic Integrity [https://policy.unt.edu/policy/06-003](https://policy.unt.edu/policy/06-003). **Plagiarism WILL result in a score of 0 for the assessment in which it occurs.**

3. Any appeals on assessment grading must be made online in Canvas (not by email) **no later than two days** after the grading is complete. This applies to examinations as well. Appeals should be as specific as possible and must contain a valid reason. Appeal such as “I think I deserve more for this assessment, or my overall grade has reduced as a result of this assessment” are **NOT valid reasons and such appeals will not be processed.** Grading is done carefully, and, in most cases, appeals will not result in a change of mark.

4. Student behavior that interferes with an instructor’s ability to conduct a class or other students’ opportunity to learn is 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 Center for Student Rights and Responsibilities 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. The Code of Student Conduct can be found at [https://policy.unt.edu/policy/07-012](https://policy.unt.edu/policy/07-012).

5. **Usage of cell phones, earphones, and other electronic devices, or recording of lectures is strictly prohibited.** Usage of laptops and tablets is permitted for class purposes, only after obtaining permission from the instructor. Usage of classrooms computers, if any, are not allowed, while the class is in session. **Any student who uses an unauthorized device will lose 1 point (out of 100) and may be asked to leave the classroom.**

**Disabilities Accommodation:**
The University of North Texas complies with Section 504 of the 1973 Rehabilitation Act and with the Americans with Disabilities Act of 1990. The University of North Texas provides academic adjustments and auxiliary aids to individuals with disabilities, as defined under the law. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring accommodation, please see the instructor and/or contact the Office of Disability Accommodation at 940-565-4323 during the first week of class.