CSCE 4205: Machine Learning

Fall 2023

Course Information & Syllabus

Class Timings: Tu/Th 11:30 PM - 12:50 PM

Location: NTDP K150

Instructor: Dr. Zeenat Tariq

Pronouns: she/her

Office: E235M Discovery Park

Office timings: Tues 10-11 AM, Wed 4 PM - 5 PM (by appointment)

Email: zeenat.tariq@unt.edu

Course Objective: Machine learning is the process of applying algorithms to learn directly from data to make predictions and decisions without being explicitly programmed in the conventional sense (e.g., programs that learn to recognize human faces, recommend music and movies, drive autonomous vehicles etc.). The course objectives are to understand machine learning algorithms and identify challenging problems on the Web, learn how to apply machine learning algorithms to these problems, and how to use the existing tools and design new ones. The course will cover both theory and practical traditional and newly developed machine learning algorithms and their Web applications. Examples of topics include a wide variety of supervised learning methods, both regression and classification, with an emphasis on those that perform well on large feature sets, Ensemble methods are used to combine independent approaches efficiently. Unsupervised and semi-supervised methods will demonstrate the power of learning from data without an explicit training target or goal. Other topics are clustering, kernel methods and Support Vector Machines, Bayesian learning methods, semi-supervised learning techniques, and feature selection, construction, and clustering. This course is designed to give a graduate-level student a thorough grounding in the methodologies, technologies, mathematics and algorithms currently needed by people who do research in machine learning or planning for data analytics roles in the industry.
**Learning outcomes:** Students in this course will learn how to apply sophisticated algorithms to large data sets, with a focus on practical application using Python based problem solving. You will be able to create models that can make automated predictions or classifications on new data or make inferences on unlabeled data to aid in understanding and future prediction models.

**Prerequisites:** Basic knowledge on probability and statistics, basic knowledge on calculus, good understanding of data structures and algorithms. Background in machine learning is not required. Familiarity with Programming (for programming assignments and activities). Understanding of basic python language is required.

**Targeted audience:** Graduate and undergraduate students from Computer Science and related areas.

**Attendance:** Attendance is essential for success in any endeavor, and it is the same for this graduate course. Consistency in attending classes is expected.

**Textbooks:**

- There is no textbook, but lectures and topic-based tutorials will be posted via Canvas throughout the semester.

**Other Recommended textbooks:**


**Course Objectives:** By the end of the course, students will be able to:

- Implement and analyze existing machine learning algorithms, including well-studied methods for classification, regression, structured prediction, clustering, and representation learning.
- Integrate multiple facets of applied machine learning in a single system: data preprocessing, learning, regularization, and model selection.
- Properly collect and organize data to extract relevant features for learning.
- Compare and contrast different paradigms for learning (supervised, unsupervised, reinforcement etc)
- Build predictive models, both regression and classification, using a variety of modeling strategies
- Using statistical learning theory to combine multiple machine learning models via ensemble methods.
- Use unsupervised learning techniques to understand high dimensional data sets.
- Using contemporary programming languages and machine learning libraries for implementing machine learning algorithms such that they can be readily applied for practical problem solving.
- Connecting concepts from probability theory with supervised learning by implementing models based on Bayes’ theorem.
- Apply advanced techniques such as reinforcement learning for robust behavior in complex, changing environments.

**Course Schedule (subject to change based on learning rate in class)**

<table>
<thead>
<tr>
<th>Week</th>
<th>Dates</th>
<th>Course/Topics</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aug 21 - 26</td>
<td>Course Overview and Introduction to Machine learning</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Aug 27 - Sep 2</td>
<td>Categories of Machine Learning, Data representation</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Sep 3 - 9</td>
<td>Computational Foundations, Data visualization and analysis</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Sep 10 - 16</td>
<td>Supervised Learning methods, Machine Learning Terminologies,</td>
<td>Project idea/Proposal</td>
</tr>
<tr>
<td>5</td>
<td>Sep 17 - 23</td>
<td>Supervised Learning Algorithms: (kNN) Nearest neighbor decision boundary</td>
<td>Quiz 1</td>
</tr>
<tr>
<td>6</td>
<td>Sep 24 - 30</td>
<td>Tree based Methods (Decision trees, Random Forest)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Oct 1 - 7</td>
<td>Support Vector Machines</td>
<td>Assignment 1</td>
</tr>
<tr>
<td>8</td>
<td>Oct 8 - 14</td>
<td>Linear and Logistic Regression</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Oct 15 - 21</td>
<td>Bayesian Networks, Naïve Bayes Classifier</td>
<td>Quiz 2</td>
</tr>
<tr>
<td>10</td>
<td>Oct 22 - 28</td>
<td>Ensemble Methods, Bagging, Boosting, Ada boosting</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Nov 5 - 11</td>
<td>Feature engineering, selection/extraction, and transformation</td>
<td>Quiz 3</td>
</tr>
</tbody>
</table>
Teaching Philosophy

I believe that the best way to really learn and understand concepts in computer science is to implement the techniques and perform the operations in an applied manner. While the process of reading examples and proofs in textbooks and lecture notes is valuable, real learning comes through one’s own efforts in debugging and solving logical and programming problems. I have a very hands-on approach to teaching, which includes coding-on-the-fly and describing how programs work, which typically requires students to be present in the lecture. I would encourage students to use this opportunity to really learn and develop the skills covered in this class that will prepare you for entering a career in industry or furthering your education.

Course Requirements

**Quizzes, Participation activities (hands on experience), Assignment, Conceptual Project Presentation, Exam**

**Participation activities:** Activities are designed to engage you in your learning, so you can begin to apply these principles in practice and tailor them to your needs. You will get hands on different concepts/examples in machine learning. Your participation in activities will be counted towards your final grade. There will be no specific due dates for the activities. However, if the activities are not submitted within four weeks. A deadline will be set for submitting all activities. The activities will be given during the class session and expected to be completed during that time. However, if some students are unable to attend the session, they can complete it remotely.

**Assignments:** There will be 2/3 assignments in the semester. The due date will be specified once the assignment is posted. Reports are to be turned in as PDF. Code is to be turned in with both Jupyter notebook and PDF form, along with any files necessary to run your assignment. Results should be presentable, with appropriate comments for someone to follow what you have done. All assignments must be turned in individually, although students are encouraged to work together extensively.

**Project Presentation (Conceptual):** After a few weeks into the course, you will be working on project (conceptual idea) and present the work by the end of semester. You are required to work
as a team. This is your opportunity to demonstrate what you have learned in a way that reaches beyond the selection of tools, data sets, and approaches demonstrated in the course. Commonly students find a unique, complex data set and associated learning problem and apply the techniques presented in the class. The goal here is to create a coherent, completed work for presentation at the end of class. Essentially ask yourself what you would want to show an employer (or brag about to others) demonstrating what you have learned in the course.

Quizzes: These quizzes are meant to focus students on the important aspects of the readings or lectures. You will be allowed to take the open book quizzes on canvas in class in person. You will be notified about the quiz one week prior taking it.

Exam: The exam is closed book and taken in class in person. Exam cannot be missed without prior arrangements or later documented proof of extenuating circumstances.

Grading

Grades are determined by a simple points system, with a total of at least 100 pts as the goal though more than 100 points are likely. The expected distribution of points is given below, with the exact scale determined by point values given for each component - this is subject to minor modification based on actual points given.

- 25% Participation activities
- 20% Assignments (2-3)
- 25% Presentation (Project Presentation)
- 15% Quizzes (3-4)
- 15% Exam
- Bonus (Possible one extra credit activity)

Grading Scale: A=90, B=80-89, C=70-79, D=60-69, F=0-59 pts. No exceptions. If class grades are low (e.g., I expect most students will end with A’s and B’s), rescaling of some of the components or extra assignment will be given to add points to the class.

Course Technology & Skills

Strongly Recommended: Laptop for in-class work. Please come to the class with a laptop that has at least two CPU-cores, 8GB memory, and 200GB hard drive. Securing and maintaining compatible hardware for in-class exercises is your responsibility. NO EXCEPTIONS.

UNT Policies

Academic Integrity Policy

Academic Integrity Standards and Consequences. According to UNT Policy 06.003, Student Academic Integrity, academic dishonesty occurs when students engage in behaviors including, but not limited to cheating, fabrication, facilitating academic dishonesty, forgery, plagiarism, and
sabotage. A finding of academic dishonesty may result in a range of academic penalties or sanctions ranging from admonition to expulsion from the University.

**ADA Policy**

UNT makes reasonable academic accommodation for students with disabilities. Students seeking accommodation must first register with the Office of Disability Accommodation (ODA) to verify their eligibility. If a disability is verified, the ODA will provide a student with an accommodation letter to be delivered to faculty to begin a private discussion regarding one’s specific course needs. Students may request accommodations at any time, however, ODA notices of 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 accommodation for every semester and must meet with each faculty member prior to implementation in each class. For additional information see the ODA website (https://disability.unt.edu/).

**Emergency Notification & Procedures**

UNT uses a system called Eagle Alert to quickly notify students with critical information in the event of an emergency (i.e., severe weather, campus closing, and health and public safety emergencies like chemical spills, fires, or violence). In the event of a university closure, please refer to Blackboard for contingency plans for covering course materials.

**Retention of Student Records**

Student records pertaining to this course are maintained in a secure location by the instructor of record. All records such as exams, answer sheets (with keys), and written papers submitted during the duration of the course are kept for at least one calendar year after course completion. Course work completed via the Blackboard online system, including grading information and comments, is also stored in a safe electronic environment for one year. Students have the right to view their individual record; however, information about student’s records will not be divulged to other individuals without proper written consent. Students are encouraged to review the Public Information Policy and the Family Educational Rights and Privacy Act (FERPA) laws and the University’s policy. See UNT Policy 10.10, Records Management and Retention for additional information.

**Acceptable Student Behavior**

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 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 classroom, labs, discussion groups, field trips, etc. Visit UNT’s Code of Student Conduct (https://deanofstudents.unt.edu/conduct) to learn more.
Access to Information - Eagle Connect

Students’ access point for business and academic services at UNT is located at: my.unt.edu. All official communication from the University will be delivered to a student’s Eagle Connect account. For more information, please visit the website that explains Eagle Connect and how to forward e-mail Eagle Connect (https://it.unt.edu/eagleconnect).

Student Evaluation Administration Dates

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 via IASystem Notification" (no-reply@iasystem.org) with the survey link. Students should look for the email in their UNT email inbox. Simply click on the link and complete the survey. Once students complete the survey they will receive a confirmation email that the survey has been submitted. For additional information, please visit the SPOT website (http://spot.unt.edu/) or email spot@unt.edu.

Getting Help

Technical Assistance

UIT Help Desk (http://www.unt.edu/helpdesk/index.htm)

Email: helpdesk@unt.edu

Phone: 940-565-2324

In Person: Sage Hall, Room 130

Walk-In Availability: 8am-9pm

Telephone Availability:

- Sunday: noon-midnight
- Monday-Thursday: 8am-midnight
- Friday: 8am-8pm
- Saturday: 9am-5pm

Laptop Checkout: 8am-7pm
Student Support Services

- Registrar (https://registrar.unt.edu/registration)
- Financial Aid (https://financialaid.unt.edu/)
- Student Legal Services (https://studentaffairs.unt.edu/student-legal-services)
- Career Center (https://studentaffairs.unt.edu/career-center)
- Multicultural Center (https://edo.unt.edu/multicultural-center)
- Counseling and Testing Services (https://studentaffairs.unt.edu/counseling-and-testing-services)
- Student Affairs Care Team (https://studentaffairs.unt.edu/care)
- Student Health and Wellness Center (https://studentaffairs.unt.edu/student-health-and-wellness-center)
- Pride Alliance (https://edo.unt.edu/pridealliance)

Academic Support Services

- Academic Resource Center (https://clear.unt.edu/canvas/student-resources)
- Academic Success Center (https://success.unt.edu/asc)
- UNT Libraries (https://library.unt.edu/)
- Writing Lab (http://writingcenter.unt.edu/)
- MathLab (https://math.unt.edu/mathlab)