Machine Learning Syllabus

CSCE 4205/5215: Machine Learning, Fall 2019 at University of North Texas

Course Calendar: go here for details on content and assignments

Instructor Contact

Name: Dr. Mark V. Albert
Office Location: Discovery Park F297A
Office Hours: Tuesday/Thursday before class (10:15-11:15am)
   Appointments through http://schedule.biomed-AI.com take priority
Email: mark.albert@unt.edu (preferred for simple questions outside of class)

TA Name: Namratha Urs
TA Email: NamrathaUrs@my.unt.edu
TA Office Hours: Tuesday/Thursday 1:00pm - 3:30pm
TA Office Location: CSE Help Lab, Room F232, or by email appointment, if you cannot make it during my office hours.

Communication Expectations: You will be expected to regularly check university email and attend class regularly. When you miss a class, you are expected to check the course calendar shortly after class to be aware of assignments, quizzes, and other materials. Questions not answered in class are best asked before or after class. For in-depth assistance on course content, you are expected to meet with the TA prior to meeting with the instructor. For quick questions, email is preferred and you can expect a response within 24 hours during the work week (M-F). For involved questions or discussions not appropriate for the TA, office hours are preferred. To guarantee priority, or for discussion outside of office hours, use the schedule link http://schedule.biomed-AI.com. It is possible through that link to arrange meetings in person as well as remotely through zoom, by phone, or skype.

Course Description

Machine learning is the process of applying algorithms to learn directly from data to make predictions and decisions without being explicitly programmed. 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.
Learning outcomes: Students in this course will learn how to apply sophisticated algorithms to large data sets, with a focus on practical application. The goal will be to create models that can make automated predictions or classifications on new data, or make inferences on unlabelled data to aid in understanding and future prediction models.

Course Structure

Time: Tuesday/Thursday 11:30am – 12:50pm
Location: Discovery Park B190
Credit hours: 3
Dates: August 27 - December 12, 2019

Tentative topics
- Machine Learning concepts
- Supervised learning, part 1
- Model validation and selection
- Supervised learning, part 2
- Feature selection and feature engineering
- Ensemble methods: bagging, boosting
- Unsupervised learning: PCA, ICA
- Clustering: K-means and DBSCAN
- Semi-supervised learning: label spreading
- Reinforcement learning
- Bayesian networks and markov models
- Deep learning

Course Prerequisites
- For CSCE 4205 specifically
  - CSCE 3110: Data Structures and Algorithms with a grade of C or better
  - Non-CSE majors must have Instructor consent by email, and Departmental Consent by filling out the Enrollment Assistance request form at www.cse.unt.edu/overrides. Please request, receive, and upload the instructor consent in the form.
- Experience with Python is beneficial as it is used extensively in the course, but significant prior programming experience with any language will be sufficient

Course Objectives

By the end of the course, students will be able to:
1. Properly collect and organize data to extract relevant features for learning
2. Build predictive models, both regression and classification, using a variety of modeling strategies
3. Use unsupervised learning techniques to understand high dimensional data sets
4. Apply advanced techniques such as reinforcement learning for robust behavior in complex, changing environments

Materials

All materials (readings, videos, tutorials, quizzes, and assignments) will be accessible online and posted on the course calendar on the respective class day at the latest. Readings will all be fairly dense, so please search for additional resources (e.g. wikipedia, coursera lectures) as needed. All attempts will be made to provide sufficient resources for everyone.

Textbook: Although no text is required, the following textbook is highly recommended:
- Additional online resources (however many other resources are available)

Course communication: We will be using the Canvas discussion board
- Feel free to use the forum to ask questions of the group, ask about partners for problem sets, or to make comments that the rest of the class might find useful.
- The forum is primarily for timely, supplementary communication; the course calendar will be the definitive source of requirements and course expectations.
- Contact the instructor if you are not available to access the discussion forum after the second week of the course.

Technical Requirements and Skills

Minimum Technology Requirements
- Computers are required for exams
  - In-class exams: You will need to bring a laptop on exam days. These will be done individually on your computer in class. You will be expected to connect to the UNT wireless network.
  - Quizzes and exams will use the Canvas quiz system
    - Canvas Technical Requirements
      (https://clear.unt.edu/supported-technologies/canvas/requirements)
• Computers are **optional during class, outside of exam times**
  ○ There will be occasional in-class activities along with lecture that are not required, but may help in understanding and applying the material
• Students will be expected to Download and install Anaconda Python version 3 and be able to open a Jupyter notebook

**Course Requirements**

**Readings, Quizzes, and Exams**

*Course Philosophy:* In this course you will be evaluated more often than other courses, with weekly quizzes, assignments, and three exams. Consider the points given to each as a guide to the effort expected. Of particular note: quizzes will be “light-weight” and you are allowed to take them at home. Given the amount of material, it is suggested that for assignments and quizzes you focus on being succinct, and for readings you focus on the main issues.

**Readings/Tutorials:** Generally, quizzes will be given based on the readings and tutorial material. As will be clear in the first few weeks, quizzes will test your knowledge on the most important aspects of the readings only.

**Canvas take-home quizzes:** These quizzes are meant to focus students on the important aspects of the readings or lectures. **You will be allowed to take these quizzes online, as many times as you would like.** All canvas quizzes will be due the last day of class, but it is suggested that you finish them in the suggested period in preparation for discussions and exams.

**Interview quizzes:** There will be two days in the semester where oral quizzes will be given in an interview format - e.g. if you don’t know the answer, work with the instructor to find it. Details will follow. The expectation is that most students will prefer these quizzes, and they provide a much clearer indication of your knowledge, are much more natural for problem solving, and help prepare you for technical interview formats.

**Competitions:** In the course there will be a Kaggle.com-style analytics competitions. Points will be distributed based on rank after the competition is finished, however the amount of points (or the variation in points across the class) will be minimal. These are primarily used as a barometer for your ability to apply these concepts quickly and efficiently.

**Exams:** Exam days are already posted and are considered fixed. Prior arrangements can potentially be made without loss of points, but have to be discussed. **Missed exams:** Exams cannot be missed without prior arrangements or later documented proof of extenuating circumstances.
Assignments

Assignments are designed to engage you in your learning, so you can begin to apply these principles in practice and tailor them to your needs.

Assignments are generally due at the end of the day one week after they are assigned, unless otherwise specified. Unlike other courses, with most assignments you are expected to post your solutions on the Trello board that is accessible to all other students, and you are free to observe other student solutions. However, no automated copy and paste is allowed, and direct handwritten or hand-typed copying without understanding and evaluation are obviously discouraged. Reports and presentation slides 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. Assignments are to be turned in individually, although students are encouraged to work together extensively.

As observed below in class grade points, these small assignments will have minimal impact on grades - they are for learning and self-evaluation rather than grading. That being said, everyone who posts a good-faith attempt on Trello for each assignment will receive one point per assignment.

Short presentations: There are far more tools and techniques than we can possibly cover, many which may be particularly relevant to your interests. Also, the ability to distill complex topics into a form useful for the audience is a critical skill to develop. Students will be expected to present one concept, tool, or technique which goes beyond what is covered in the course. Each presentation is to be 5-10 minutes on the whiteboard or with slides. These presentations will be concentrated on 1-2 days in the semester. Only a basic familiarity is expected of students observing the presentations.

Projects: After a few weeks into the course you will select among a small number of collaborative projects. Project proposals, progress reports, and final reviews will be part of the process. You are required to work in groups, as this is part of a full and complete education. All people in the group are expected to contribute. 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 project 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.
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 assignment, quiz, or exam - this is subject to minor modification based on actual points given. Note, due to the nature of the course, exams and quizzes are a significant means of establishing your final grade, so please complete the assignments in a timely way and study appropriately prior to each quiz and exam.

- Assignments and take-home quizzes: 15 pts
  - Canvas take-home quizzes: 5pts (~0.5 pts each)
  - Assignments: 10 pts (~1 pt each)
- Presentation and Project: 25 pts
  - Short presentations: 5 pts
  - Projects: 20 pts
    - Brainstorming: 1 pt
    - Proposal: 2 pts
    - Update: 2 pt
    - Report: 10 pts
    - Project presentation: 5 pts
- Interview quiz and competitions: 20 pts
  - Interview quiz: 10 pts (~2 pts each)
  - Competitions: 10 pts (5 pts each)
- Exams: 45 pts
  - Exam I: 10 pts
  - Exam II: 15 pts
  - Final Exam: 20 pts

Grading Scale: A=90, B=80-89.9, C=70-79.9, D=60-69.9, F=0-59.9 pts. No exceptions. If class grades are low (e.g. I expect the vast majority of students will end with A’s and B’s), extra quizzes or assignments will be given to add points to the class. (Note, these grades are based on points and not percentages, so if 120 points are given, you only need 93/120 for an A!)

CSCE 4205 vs 5215 grading: In line with the added expectations for the graduate version of the course, students enrolled in CSCE 4205 will receive 5 additional bonus points.

Course Evaluation

Student Perceptions of Teaching (SPOT) is the student evaluation system for UNT and allows students the ability to confidentially provide constructive feedback to their instructor and
department to improve the quality of student experiences in the course. Spot evaluations will be available November 18 - December 5th, 2019

Course Policies

Examination Policy
Exams will be on the computer using the Canvas quiz system. You need to bring a laptop on the appropriate exam days. **Exams are “closed book”** - no use of materials outside the canvas exam system unless otherwise specified in advance. Quizzes and exams must be taken in the classroom unless special accommodations have been made through the Office of Disability Accommodation (ODA). Another other accommodations must be given by prior arrangement with the instructor, otherwise documentation proving an extenuating circumstance will have to be provided after the missed exam. Time will be limited, and all work will be individual. You are strongly encouraged to attempt to solve the tasks iteratively and incrementally - write code that works first, but works poorly, and improve from there, rather than write perfect code top to bottom. Exams will focus on the most recent material but are expected to be cumulative in scope.

Technical errors during exams
If during an online quiz or exam there is a technical error which affects your ability to complete the assignment, you are immediately to let the quiz or exam proctor know and the instructor will discuss ways to allow you to resume the test without giving an unfair advantage. In the event of any unexpected server outage or any unusual technical difficulty which prevents students from completing a time sensitive assessment activity, the instructor will extend the time windows and provide an appropriate accommodation based on the situation. Students should immediately report any problems to the instructor.

Late Policy
When assignments and project work are turned in after the due date, this places an undue burden on the instructor and TA, especially when this policy is abused. As a compromise, if the assignment or project work is turned in **prior to grading** there will be no reduction in points, however, grading can occur any time after the due date (including the following morning!). A request for missing submissions may be sent at the discretion of the TA or instructor, but is not guaranteed, and a reduction of points may also occur in a way that is consistent for the rest of the class. Given the frequency of assignments and quizzes, and this flexible late policy, there may be a significant lag between submission and entry into the grade book for already-documented small-point assignments.

Attendance Policy
You are expected to attend lectures and to complete all readings, however, this course does not use participation points and there is no penalty for missing days without exams, or group project efforts. There is no need to let the instructor know you have missed a class, however, you are responsible for keeping up with the material covered in the class if you are not present. If a
If a class is missed, you are expected to proactively reach out to classmates, the TA, or the instructor if there are any questions.

Individual attendance is not required except on group project and exam days. There is no direct participation grading, but in the past there has been a strong correlation between engagement and accomplishment in courses - especially for those that are struggling with the material. Feel free to prioritize your time, but prioritize wisely.

Syllabus Change Policy

Any substantial changes to the syllabus after the first week will be highlighted in red on the online platform. Approximate point values are expected to vary but will be fixed when the assignment or exam is given.

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)