Machine Learning

Dr. Dan DeBlasio
Department of Computer Science
University of Texas at El Paso

CS 4361/5361 – Fall 2022

Revised: August 17, 2022
(change since start of classes marked in orange)

Course Description: Machine Learning studies the development of programs that can improve in the performance of a task with experience. For many difficult problems, such as speech understanding, image classification, and text analysis, solutions based on machine learning outperform all others proposed to date.

In this course we will study several of the most commonly used machine learning algorithms, their application to problems in several areas of interest, and their quantitative evaluation. We will also discuss current research issues in machine learning. Each student will do a research project related to a problem of their interest.

Prerequisite: CS 2302, MATH 3323 (CS 4361 only) , STAT 3320 (CS 4361 only) or instructor approval.

Contents

1 Logistics

Synchronous course session times and locations:

- TR 4:30pm-5:50p — CCSB G.0208

Textbook: There will be no textbook. We will use parts of online books and other materials.

2 Instructional Staff

Instructors

Dr. Dan DeBlasio

email:	dfdeblasio@utep.edu
office:	CCSB 3.1008
office hours:	tbd
appointments:	calendly.deblasiolab.org

3 Expectations

Students are expected to consult their emails and blackboard messages *at least* twice a week, and to answer these as relevant.

As a courtesy to their peers, students should be on time for all scheduled sessions and attend the entire session.

Professionalism: Students should be professional in their communications, as the context permits. Emails should contains subjects, the recipients should be addressed (i.e. "Hello Dr. DeBlasio, ..."), and the email should be signed with your name. Real-time online communication (i.e. MS Teams), while less formal, should still be professional.

4 Grading

Grades are communicated to students in a timely manner. It is the students' responsibility to keep track of their grades by compiling the grades they receive.

The approximate percentages are as follows:

4361	5361	
5%	5%	Lab assignments
15%	15%	Participation, homework, quizzes
50%	40%	Tests
10%	20%	Final project

> **Graduate vs. Undergraduate**
>
> As reflected by the relative weights, the expectations for the final project are much higher for students taking CS 5361

Conversion to a letter grade assignment

90%	or higher is guaranteed an A
80%	or higher is guaranteed a B
70%	or higher is guaranteed a C
60%	or higher is guaranteed a D
	all lower grades are an F

These minimums may be lowered without notice but will not be raised.

4.1 Lab assignments

Lab grades will be reduced by a factor of 25% for each working day or fraction they are late.

4.2 Participation, homework, quizzes

Regular attendance is essential and expected. Due to the high emphasis of group discussion and dialogue all students are discouraged from missing classes. Missed course meetings will be noted and chronic absences may impact the students grade if not discussed with the course instructor.

To asses the progress of the students in the class, quizzes and homework will be assigned without notice throughout the semester. Quizzes will be graded on completion not on correctness and are used by the instructor. Homework will be assigned and graded for completeness and *may* include a coding component.

4.3 Tests

All exams will be administered during class.

Students will be randomly selected to explain, during a demo session, their exam solutions. This examination will be used to determine the student's exam grade regardless of the received number of points.

5 Standing in the course

Special Assignments: If deemed necessary, special assignments will be given to students to ensure that said students remain in the class and be successful. These will be designed to help students grow into the course and develop the necessary skills. It is important that students feel free to ask their instructor about any such opportunity as well so that a special plan of development for CS2401 be tailored to them.

Standing in the Course: Students will have access to their grades for all assignments so that they can self-monitor their standing and progress. However, it is also completely fine for any student to come and talk to their instructor about their standing and work together to make sure the student is as successful as can be.

Dropping the Course: Every semester, some students drop the course. We, instructors, completely understand and respect that. We only hereby ask students to inform us, ideally before, but in the worst-case right after, of their intention to drop the course. This is

really important for us as it possibly informs us of ways in which to better serve our students.

6 Special notices for COVID-19

Please stay home if you have been diagnosed with COVID-19 or are experiencing COVID-19 symptoms. If you are feeling unwell, please let me know as soon as possible, so that we can work on appropriate accommodations. If you have tested positive for COVID-19, you are encouraged to report your results to covidaction@utep.edu, so that the Dean of Students Office can provide you with support and help with communication with your professors. The Student Health Center is equipped to provide COVID-19 testing.

The Center for Disease Control and Prevention recommends that people in areas of substantial or high COVID-19 transmission wear face masks when indoors in groups of people. The best way that Miners can take care of Miners is to get the vaccine. If you still need the vaccine, it is widely available in the El Paso area, and will be available at no charge on campus during the first week of classes. For more information about the current rates, testing, and vaccinations, please visit epstrong.org.

> **Masks in the classroom**
>
> **To this end, until further notice the instrucional team and students are in the strongest terms possible encouraged to wear a mask at all times in the classroom.** If this is not possible, or you don't feel this is something that you want to do you're encouraged to transfer to another section of the course.

7 Resources

Special Accommodations: If you have a disability and need classroom accommodations, please contact the Center for Accommodations and Support Services (CASS) at 747-5148 or by email to cass@utep.edu, or visit their office located in UTEP Union East, Room 106. For additional information, please visit the CASS website at `www.sa.utep.edu/cass`. CASS' staff are the only individuals who can validate and if need be, authorize accommodations for students with disabilities.

Scholastic Dishonesty: Any student who commits an act of scholastic dishonesty is subject to discipline. Scholastic dishonesty includes, but not limited to cheating, plagiarism, collusion, and submission for credit of any work or materials that are attributable to another person.

Cheating is:

- Copying from the test paper of another student

- Communicating with another student during a test to be taken individually
- Giving or seeking aid from another student during a test to be taken individually
- Possession and/or use of unauthorized materials during tests (i.e. crib notes, class notes, books, etc.)
- Substituting for another person to take a test
- Falsifying research data, reports, academic work offered for credit

Plagiarism is:

- Using someone's work in your assignments without the proper citations
- Submitting the same paper or assignment from a different course, without direct permission of instructors

  To avoid plagiarism, see:
  https://www.utep.edu/student-affairs/osccr/_Files/docs/Avoiding-Plagiarism.pdf

Collusion is:

- Unauthorized collaboration with another person in preparing academic assignments

> **Important!**
>
> When in doubt on any of the above, please contact your instructor to check if you are following authorized procedure. Also, please check the UTEP's Handbook of Operating Procedures at: hoop.utep.edu.

8 Tentative List of Topics to be Covered

1. Introduction
   a. Machine Learning – what, why and how.
   b. Machine Learning tasks: classification, regression, and reinforcement learning
   c. Basic tools
2. Evaluating Performance
   a. Metrics
   b. Training and test sets
   c. Cross-validation
      i. N-fold cross validation
      ii. Leave-one-out cross validation
3. Supervised Learning algorithms
   a. K-nearest neighbors
   b. Naïve Bayes classifier
   c. Logistic regression
   d. Decision and regression trees
   e. Random forests
   f. Evolutionary algorithms
   g. Support vector machines

---

       h. Neural networks
4. Unsupervised Learning and preprocessing
       a. Principal component analysis
       b. K-means
5. Applications
       a. Computational Biology
       b. Natural language processing
       c. Computer vision