

CPS 5320 Syllabus, Computational Science, UTEP

Course #:	CPS 5320
Course title:	Advanced Scientific Computing
Credit hours:	3
Term:	Fall 2020
Time & location:	12:00 -1:20pm MW, Miners Hall 201
Prerequisites:	Calculus III (MATH 2313), Matrix Algebra (MATH 3323), and Introduction to Computational Science (CPS 5401) with a grade of B or better; or permission of the instructor.
Course fee:	None
Instructors:	Suneel Babu Chatla
Office location:	Bell Hall Room 227
Contact information:	Office phone: 915-747-6764 Email: sbchatla@utep.edu
Office hours:	2:00-3:00 pm MW or by appointment.
Course Materials	<p><u>Textbook References:</u></p> <ul style="list-style-type: none">● James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). <i>An introduction to statistical learning</i> (Vol. 112, p. 18). New York: springer. (available online)● Bishop, C. M. (2006). <i>Pattern recognition and machine learning</i>. springer.● Shmueli, G., Bruce, P. C., Yahav, I., Patel, N. R., & Lichtendahl Jr, K. C. (2017). <i>Data mining for business analytics: concepts, techniques, and applications in R</i>. John Wiley & Sons. <p><u>Software:</u> In this course, I primarily use R software. Of course, you are free to use any other software such as Python or Matlab. Both R and Python are open source and can be easily installed on any of the popular operating systems.</p>

CPS 5320 Syllabus, Computational Science, UTEP

Course Description

This course aims to provide an introduction to the statistical learning and computing. With the recent advances in computing power, many statistical learning methods that are highly computational are developed. These methods have huge potential in solving or providing better understanding of the real life problems. At the end of this course you should have a basic understanding of how all these methods work and be able to apply them in the fields you are working on. With the explosion of “Big Data” or “Data Science”, statistical learning has become a very hot field in many scientific areas and people with these skills are in high demand.

To this end, approximately one third of the class time is dedicated to in class labs where we work on the methods we have covered until then. These labs will ensure that every student has a full understanding of the practical, as well as the theoretical, aspects of each method.

Grading

Your grade for the course will be based on the following:

- 40% homework
- 25% midterm and 25% final project
- 10% Kaggle Competition

Attendance Policy

Lecture and lab attendance is required and noted at the beginning of class; more than a total of TWO unexcused absences will result in an instructor-initiated drop or final grade reduction. Your academic/research advisor will be consulted before final action is decided and taken.

Accommodations for Students with Disabilities

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.

CPS 5320 Syllabus, Computational Science, UTEP

Academic Honesty Policy

Make sure you understand the UTEP academic honesty policy. Students are encouraged to share ideas, but you must do your own homework and you must write your own code for the projects (you may copy code that is on the course website). If homework or program code is suspected of being duplicated or copied, you will receive an incomplete for the assignment, and your case will be referred to the Dean of Students for adjudication. If the instructor has reason to believe that you have cheated on a quiz or exam, your case will be referred to the Dean of Students for adjudication.

Course Format and Participation

The lecture portion of the class will consist of short lectures interspersed with hands-on interactive activities. Lab assignments will reinforce the lecture material. The lecture and lab exercises will make use of CS Department Linux lab machines, the UTEP Research Cloud, and the Stampede Supercomputer at Texas Advanced Computing Center. Students should be able to login to these resources remotely from a home or office computer. All students should bring a laptop computer to class with which to login to the remote resources. (Please let the instructor know if you do not have a laptop you can bring to class).

Course Topics

The following is a list of topics to be discussed. The exact schedule may vary depending on previous background of the class participants.

1. Introduction
 - What is statistical learning?
 - Probability basics
 - Introduction to R
2. Data Pre-processing
 - Data transformations
 - Missing values
3. Assessing Model Accuracy
 - The Bias-Variance Trade-Off
 - Missing values
 - The Classification Settings
4. Linear Regression
 - Simple linear regression

CPS 5320 Syllabus, Computational Science, UTEP

- Multiple linear regression
5. Classification
 - Linear discriminant analysis
 - Logistic regression
 - Neural networks
 6. Resampling methods
 - Cross-Validation
 - Bootstrap
 7. Model Selection and Regularization
 - Stepwise Selection
 - Shrinkage
 - Dimension reduction
 - High-Dimensional Data
 8. Curve estimation
 - Polynomial regression
 - Splines
 - Generalized additive models
 9. Tree-Based Methods
 - Regression Trees
 - Classification Trees
 - Random Forests
 - Boosting
 10. Support Vector Machines
 11. Deep Learning

Term project

The term project will consist of:

1. development and solution of a statistical model for a physical problem not discussed in class nor assigned for homework
2. a report describing the background for the model, model definition, statistical methods, results, and validation
3. a presentation describing and demonstrating your model,

CPS 5320 Syllabus, Computational Science, UTEP

The specific problem can be of your choosing but you must have your topic per-approved by the instructor. You may work individually on the final project or in teams of up to two people. In case of group work, you must clearly document the contributions of each team member and carry out the amount and difficulty of the work proportional to the size of your team.

Military Statement: If you are a military student with the potential of being called to military service and/or training during the course of the semester, you are encouraged to contact as soon as possible.