460J (17756, 17790, 17795, 17800)
Data Science Lab
Fall 2022

Instructors:
Alex Dimakis (dimakis@austin.utexas.edu). Office: EER 6.816
Office hours: 2-3pm Thursdays in EER Office.

TAs: Weiran Huang <huangweiran@utexas.edu> Office hours: TBD.
Ritika Mangla <ritikamangla@utmail.utexas.edu> Office hours: TBD.

Lectures: T.Th. 3.30-5pm, ECJ 1.202.

Course Description: The emerging field of data analytics and, more broadly, data science, is transforming engineering, healthcare, scientific discovery and many industries ranging from Agriculture to Telecommunications. In this class we are going to discuss how to use data to build models to perform prediction and inference. Topics: Predictive modeling. Regression and Classification. Data cleaning and preprocessing. Feature engineering. Unsupervised methods. Principal Component Analysis. Data clustering. Model selection and feature selection. Entropy and Information theory. Neural Networks and Deep Learning. Machine learning for signals and time-series data.

Data Science Lab vs Data Science Principles: Necessarily, this class has significant overlap with the Data Science Principles course, as we have worked hard to ensure that they can be taken in any order. Both classes cover basics like regression, classification, decision trees, bagging and boosting. However, the lab expands much more on model tuning, performance, practical issues and challenges. On the other hand the principles course develops the theory, analysis and derivations. Also, the lab covers neural networks in much more detail compared to principles.

Requirements: The class will have a significant hands-on component on working with real data, performing modeling and prediction. Laptops with wifi access, Python, Numpy, Pandas and Scikit installed are required. Additional tools will be discussed when introduced.

Course Material:

Online resources will be posted on Canvas, and also on the course web page:
http://users.ece.utexas.edu/~dimakis/DataScienceLab.html

Additionally we recommend:
ISL Book: An Introduction to Statistical Modeling (G. James et al.),
And other materials we will refer to in each class as needed.

Deep learning by Goodfellow, Bengio and Courville
https://www.deeplearningbook.org/
And other materials we will refer to in each class as needed.

Grading:

- Labs reports and Quizzes: 35%
- Kaggle competition 20%.
  Opens Monday October 17th, Closes Sun. October 23st.
- In-class midterm: 15% Tuesday November 8th
- Final Project: 30%. **Teams of 6**.
    - The project includes a report and presentation
    - 40% of the project grade will be from the presentation.
    - Presentations will be pre-recorded videos that the teams make (up to 10 minutes) due on the last day of classes. The instructors can reach out to a team for questions and clarifications.
    - 60% of the project grade will be from the report, also due on the last day of classes.
      Note: The last day of classes is set by the UT academic calendar. It does not refer to the last lecture of this particular class.
  - Project proposal -- who's on your team, what you want to do and with which data set(s) you plan to use.
  - There will be no final exam.

Class Policy Rules

1. This is a very large class. Unfortunately, we cannot modify pre-requisite requirements on a student by student basis.
2. In each lecture you are required to bring a laptop or another device you use to take canvas quizzes with. *In each lecture we may ask you to take an in-class canvas quiz*.
3. Lab reports will be completed in **teams of 3**. Teams should be in the same lab section.
4. Late homeworks and lab reports: You can be late up to 2 hours. After that, no late homeworks and lab reports will be accepted. To compensate for late homeworks, ***we will drop your lowest lab score****. Other than this, we cannot make any exceptions to the late policy*.
5. Sign up for one lab section. Each lab section has a specific number of seats. You must have a partner in the same lab section for the lab reports.
6. Attending your lab section is not mandatory. You have to obviously turn in your lab report to get graded. Do not attend another lab section since each lab room has a specific number of seats.
7. Write all FULL names and UT EIDs on each lab report and project report.
8. Discussion of homework and lab reports is encouraged. Please be absolutely sure to submit your own solutions that you type yourself.
9. When submitting a lab report or a project report on canvas, EVERY participating student in that report should upload the materials on canvas, not only one of them.
10. Your project report submission will be a blog post. If you have large data sets, don't worry – we will help find the resources to help you post this. For some examples, of past projects, you can see some reports are here: http://users.ece.utexas.edu/~dimakis/DataScienceLab.html see 'Some Final Projects from Spring 2017'.
11. Do not leave cloud-running instances on, if you are not using them. You will lose your credits quickly. Some students in the past created thousands of dollars of debt by forgetting their AWS instances running.
12. **Graduate Students:** If you are a graduate student, we prefer that you partner with other graduate students for the labs, and also for the final projects.

University Honor Code

“The core values of The University of Texas at Austin are learning, discovery, freedom, leadership, individual opportunity, and responsibility. Each member of the University is expected to uphold these values through integrity, honesty, trust, fairness, and respect toward peers and community.”

College of Engineering Drop/Add Policy
The Dean must approve adding or dropping courses after the fourth class day of the semester.

Students with Disabilities
UT provides upon request appropriate academic accommodations for qualified students with disabilities. Please contact the Office of Dean of Students at 471-6259 or ssd@uts.cc.utexas.edu.

Emergency Preparedness
Every member of the university community must take appropriate and deliberate action when an emergency strikes a building, a portion of the campus, or entire campus community. Emergency preparedness means we are all ready to act for our own safety and the safety of others during a crisis.
Students requiring assistance in evacuation must inform the instructor in writing of their needs during the first week of class. This information must then be provided to the Fire Prevention Services office by fax (512-232-2759), with "Attn. Mr. Roosevelt Easley" written in the subject line.
You may want to bookmark the emergency Web site http://www.utexas.edu/emergency/ because it is updated with information during actual emergencies or campus closures. The university collects cell phone numbers from members of the campus community for emergency text messages. You can sign up for campus text alerts online. If you would like more information regarding emergency preparedness, visit http://www.utexas.edu/safety/preparedness

COVID-19 Pandemic

The COVID pandemic is creating a rapidly changing landscape. We will be covering part of the lectures online as needed by the pandemic and as dedicated by UT rules and guidelines. Please stay updated on UT COVID policies.

COVID-19 Guidance
To help preserve our in-person learning environment, the university recommends the following.
• Adhere to university mask guidance.
https://utexas.app.box.com/s/ymob0b4vimv4j9gnhskpsqywwadk3f10

• Vaccinations are widely available, free and not billed to health insurance. The vaccine will help protect against the transmission of the virus to others and reduce serious symptoms in those who are vaccinated.
https://uthealthaustin.org/patient-resources/covid-19-updates/covid-19-vaccination

• Proactive Community Testing remains an important part of the universitys efforts to protect our community. Tests are fast and free.
https://healthyhorns.utexas.edu/coronavirus_proactive_testing.html

• Visit Protect Texas Together for more information
https://protect.utexas.edu/

For virtual zoom lectures, please attend in person since the instructors will do their best to keep them interactive. Also please turn on your cameras if you want to-- this makes it easier for us to connect with the students and understand if there are questions.

The COVID-19 pandemic has been disrupting our health and well-being for several months. More than 1.7% of the US population has tested positive, including friends and family in some cases. Its cumulative effect has taken a serious toll on our physical and mental health. Many of our friends and family have lost their jobs or been furloughed. Many of us are living in an environment that may not have a quiet place to study or adequate Internet access. Please let me know if you face difficulties this semester in accessing course resources or completing work. The University provides student emergency assistance through Texas One Stop (US citizens), Texas Global (International), Student Emergency Fund (All Students).

Video Streaming and Recording

Video streaming and recording of class activities are reserved only for students and TAs in this class for educational purposes and are protected by FERPA laws if any students are identifiable in the video. Video streams and recordings should not be shared outside

the class in any form. Students violating this university policy could face misconduct proceedings.