

Department of Electrical Engineering
University of Southern California

**EE 565 – INFORMATION THEORY and its APPLICATION TO (BIG) DATA
SCIENCES Spring 2024**

Instructor: Urbashi Mitra, Professor
536 EEB, 213 740 4667, ubli@usc.edu Office Hours: TBD

Teaching Assistant: TBD

Course Web Page: DEN Blackboard www.uscdcn.net
Contains homework, solutions, and relevant handouts. Course announcements, homework hints and modifications will be posted on this page – please check it regularly.

Lectures: MW 10:00 am - 11:50pm, TBD

Course Objectives: Information theory answers two fundamental questions in communication and compression theories: What is the ultimate data compression (answer: the entropy H), and what is the ultimate transmission rate of communication (answer: the channel capacity C). This course covers fundamental theories and practical algorithms for both data compression and reliable communication. At the completion of the subject, students will have mastered basic concepts and tools in information theory, and will be able to analyze a wide range of problems in data compression and data communication over noisy channels. They will also be able to design algorithms for data compression, and design error correction codes for data communication. Topics include: entropy and other information measures; variable and fixed-length lossless and lossy source codes; universal compression; applications to text and multimedia compression; channel capacity and the channel coding theorem; error-correcting codes and real channels; and statistics and data science. **As many machine learning and neural network algorithms employ information theoretic measures (e.g. cross-entropy, Kullback-Leibler divergence, mutual information etc.) as key objective functions, the Spring 2024 offering of 565 will specifically look into data-driven driven strategies for estimating these quantities. Direct estimation of probability mass or density functions from data and then computing measures usually does not work well.**

Prerequisites: Probability theory and random variables, moments, transformations of random variables, *etc.* (EE 503).

Other Requirements: Basic computer skills (*i.e.* programming, plotting, random variable generation, familiarity with Matlab is helpful although not necessary), knowledge of convex functions and their properties, simple optimization, stationary points of functions, limits of sequences and series, *etc.*

Text: Elements of Information Theory, by T. M. Cover and J. A. Thomas, Wiley, 2nd edition, 2006.

A Brief Introduction to Machine Learning for Engineers, by O. Simeone, Foundations & Trends in Signal Processing, NOW Publishers, 2018 (will be posted in Blackboard)

Grading: (tentative) 20% Homework
25% Midterm
30% Final
25% Projects (8% and 17%) (if there are three projects the total distribution across all graded items will be adjusted)
Final grades will be assigned by a combination of student score distribution (curve) and the discretion of the instructor.

Exams: Midterm (tentative) Monday, March 4, 2024, 10:00 am - 11:50pm
Final Monday, May 6, 2024, 8:00-10:00am

Office Hours: TBD

Use of email to set up appointments encouraged: ubli@usc.edu. Attending office hours in person is encouraged.

Late Policy: Homework is due at 11:59pm on Wednesdays. No late homework will be accepted. A late assignment results in a zero grade.

Make-up Material: Homework assignment dates are non-negotiable. Your lowest homework score will be thrown out before computing final grades. No make-up exams will be given. In the case of a required business trip or a medical emergency, a signed letter from your supervisor or doctor is required. This letter must include the telephone number of your doctor or supervisor.

Grade Adjustment: If you dispute any scoring of a problem on an exam or homework set, you have **one week** from the date that the graded paper is **returned** to request a change in the grade. After this time, no further alterations will be considered. All requests for a change in grade must be submitted in writing to me.

Attendance: Lecture attendance is encouraged; many examples and applications not in the text will be covered in the lectures. The student is responsible for all assignments, changes of assignments, announcements, lecture notes *etc.* All such changes should be posted on the course web-site.

- References:**
1. Information Theory Lecture Notes, S. Moser, 2019.
 2. Advanced Topics in Information Theory Lecture Notes, S. Moser, 2022.
 3. Information Theory, Inference, and Learning Algorithms, D. Mackay, Cambridge University Press, 2003.
 4. A First Course in Information Theory, R. Yeung, Springer 2002.
 5. Principles of Digital Communication, R. Gallager, Cambridge University Press, 2008.

Topics (sections refer to Cover & Thomas)

1. Basic information measures (2.1-2.3)
2. Properties of information measures; typicality (2.4-2.9,3.1-3.3)
3. Sufficient statistics *information bottleneck principle (notes)*

4. Maximum likelihood, supervised learning, connection to information theoretic measures (notes)
5. *Computation of information measures from data*
6. **First project - data driven entropy estimation**
7. Lossless compression, Huffman coding (5.1-5.8)
8. Universal compression (and learning); entropy rate; compression of stationary sources (13.1-13.2)
9. Algorithms: arithmetic coding; Lempel-Ziv; (13.3-13.5)
10. **Second project - compression**
11. Binning; Slepian-Wolf coding (15.4,15.8)
12. Point-to-point communications; channel capacity (7.1-7.7)
13. Channel coding theorem for DMCs (7.8-7.10)
14. Examples of error-correcting codes; decoding algorithms (7.11)
15. Differential entropy; channel coding theorem for Gaussian channels (8.1-8.6,9.1-9.2); *computing Kullback-Leibler divergences of Gaussian mixtures from data*
16. Bandlimited channels; parallel channels (9.3-9.4)
17. Rate-Distortion Theory (10.1-10.5,10.8) ; *revisiting the information bottleneck principle*
18. Information theory and statistics (11.1, 11.2, 11.5, 11.7-11.9)
19. **Third project (?): classification/statistics project**
20. Feedback (7.12-7.13, 9.6)
21. Multiple-access channels (15.3)
22. Broadcast channels (15.6)

- Suggestions:**
1. Remember the big picture.
 2. Read the book and supplementary sources.
 3. Prepare your own summaries from texts and notes.
 4. Work as many problems as you can.

Academic Conduct: Plagiarism (presenting someone else's ideas as your own, either verbatim or recast in your own words) is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Section 11, Behavior Violating University Standards . Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, . Note that the posting of homework and exam solutions in a public forum also constitutes academic misconduct. Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the Office of Equity and Diversity (see below) or to the Department of Public Safety (see below). Anyone in the university community such as a friend, classmate, advisor, or faculty member can help initiate the report, or can initiate the report on behalf of another person. The Center for Women and Men provides 24/7 confidential support, and the sexual assault resource center webpage sarc@usc.edu describes reporting options and other resources.

Disabilities: The Office of Disability Services and Programs provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, USC Emergency Information <http://emergency.usc.edu/> will provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology.

Support Systems: *Student Counseling Services (SCS)* - (213) 740-7711 – 24/7 on call Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

<https://engemannshc.usc.edu/counseling/>

National Suicide Prevention Lifeline - 1-800-273-8255 Provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week. <http://www.suicidepreventionlifeline.org>

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-4900 - 24/7 on call Free and confidential therapy services, workshops, and training for situations related to gender-based harm. <https://engemannshc.usc.edu/rsvp/>

Sexual Assault Resource Center For more information about how to get help or help a survivor, rights, reporting options, and additional resources, visit the website: <http://sarc.usc.edu/>

Office of Equity and Diversity (OED)/Title IX Compliance – (213) 740-5086 Works with faculty, staff, visitors, applicants, and students around issues of protected class. <https://equity.usc.edu/>

Bias Assessment Response and Support Incidents of bias, hate crimes and microaggressions need to be reported allowing for appropriate investigation and response. <https://studentaffairs.usc.edu/bias-assessment-response-support/>

Student Support and Advocacy – (213) 821-4710 Assists students and families in resolving complex issues adversely affecting their success as a student EX: personal, financial, and academic. <https://studentaffairs.usc.edu/ssa/>

Diversity at USC Information on events, programs and training, the Diversity Task Force (including representatives for each school), chronology, participation, and various resources for students. <https://diversity.usc.edu/>

American Language Institute <http://dornsife.usc.edu/ali>, sponsors courses and workshops specifically for international graduate students for whom English is not their primary language.

USC Emergency Information Provides safety and other updates, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible, <http://emergency.usc.edu>

USC Department of Public Safety – 213-740-4321 (UPC) and 323-442-1000 (HSC) for 24-hour emergency assistance or to report a crime. Provides overall safety to USC community. <http://dps.usc.edu>