

## Math 547, Statistical Learning Theory, Fall 2021

**Exterior Course Website:** <http://www.stevenheilman.org/~heilman/547f21.html>

**Prerequisite:** Working knowledge (graduate or advanced undergraduate level) of the following: Probability Theory (Math 407/505a, 507a recommended) Real Analysis (Math 425a/b, Math 525a recommended) and Linear Algebra (Math 471).

**Course Content:** Binary classification, empirical risk minimization, support vector machines, voting algorithms and AdaBoost, Vapnik-Chervonenkis theory, concentration of measure inequalities, sparse recovery problems, high-dimensional convex geometry.

**Lecture Meeting Time/Location:** Mondays, Wednesdays, and Fridays, 10AM-1050AM, KAP 140

**Instructor:** Steven Heilman, [stevenmheilman@gmail.com](mailto:stevenmheilman@gmail.com)

**Office Hours:** Tuesdays, 10AM-12PM noon, on zoom [link posted on blackboard]

**TA:** Jihoon Sohn, [jihoon.sohn@usc.edu](mailto:jihoon.sohn@usc.edu)

**TA Office Hours:**

**Textbooks, etc.:** There is no official textbook, but the following might be useful references Kearns and Vazirani, An Introduction to Computational Learning Theory. I will be following this book the most in the first half of the course.

Shalev-Schwarz and Ben-David, Understanding Machine Learning: From Theory to Algorithms. (A link is available [here](#).)

Giné and Nickl, Mathematical Foundations of Infinite-Dimensional Statistical Models. This encyclopedic book has a lot of nice mathematical tools in it. Chapters 2 and 3 are most relevant.

Vershynin, High-Dimensional Probability: An Introduction with Applications to Data Science. (A link is available [here](#).) Tools and applications. Ledoux, Concentration of Measure. Another book with useful mathematical tools. Most chapters here are relevant.

Boucheron, Lugosi, Massart, Concentration Inequalities. An encyclopedic book with some overlapping content with Ledoux's book; chapters 12 and 13 are most relevant..

Mossel's Course Webpage on Deep Learning, available [here](#), based on the book Deep Learning, by Goodfellow, Bengio and Courville, available [here](#). **This course has no midterm exams and no final exam.** Instead, we will have homeworks and a final project.

**Project Abstract Due:** Friday, September 24, 5PM PST (via blackboard)

**Progress Report Due:** Friday, October 29, 5PM PST (via blackboard)

**Final Project Presentations:** Last two or three weeks of class, in class, schedule TBD

**Final Report Due:** Saturday, December 4, 5PM PST (via blackboard)

This course has no midterm exams and no final exam. The final project is a replacement for the exams.

### Email Policy:

- My email address for this course is [stevenmheilman@gmail.com](mailto:stevenmheilman@gmail.com).
- It is your responsibility to make sure you are receiving emails from [stevenmheilman@gmail.com](mailto:stevenmheilman@gmail.com), and they are not being sent to your spam folder.
- Do NOT email me with questions that can be answered from this document.

**Final Project Guidelines:** The final project is an opportunity to do a bit of research on a subject of your choice related to statistical learning theory. A project could begin with an interesting question or a well-known problem, and perhaps lead to investigating or implementing various algorithms, conducting an empirical analysis, etc.

Along the way, you will review relevant literature, identify appropriate data sources, select appropriate means of evaluation, and either develop novel methodology for your problem or deploy and comprehensively evaluate existing methodology for your new application.

The goal is to say something interesting about a problem in statistical learning theory. You could perhaps develop new methodology for an existing problem or application that has no fully satisfactory solution. You could alternatively tackle a new problem or application with existing methodology; in this case, you should identify one or more questions without satisfactory answers in your chosen domain and explore how the methodology can help you answer those questions. You may draw inspiration from particular datasets, but your focus should rest not on the data itself but rather on the questions about the world that you can answer with that data.

While a substantial theoretical component is not required for this project, (i.e. you do not necessarily have to prove new theorems), you should be able to prove a few things or at very least review or streamline some existing proofs in your investigation.

You may work alone or in a group of two; the standards for a group project will be twice as high. In certain cases I might approve a group of three, but this is unlikely.

We strongly encourage you to come to office hours to discuss your project ideas, progress, and difficulties.

**Final Project Milestones:** In all cases, please use LaTeX.

I: Project Proposal. By this first milestone, you should have selected a question or problem of interest, identified relevant data sources, begun exploring the literature surrounding the question, and discussed your ideas with the course staff. Your project proposal deliverable is a 1/2 - 1 page report describing the question or problem you intend to tackle, why this question is important or interesting, prior work on this problem, what data you intend to use in your analyses, and the principal challenges that you anticipate.

If you would like to receive feedback about particular aspects of your proposal, please indicate this in your submission.

I can try to help in problem selection. Ideally, the problem should be something you are very interested in. As such, it might be helpful to first tell me about your interests (maybe after class or in office hours), and we can try to think of something to work on. Selecting problems to work on is a difficult skill that takes years to develop, so it would be nice if you find a project idea on your own, but I expect everyone will need at least a little help in their choice. I know some things about some fields but I don't know everything about every field, so I might not be so helpful with certain projects outside my own background, but I can learn a bit myself to help you along if your interests are outside my knowledge.

II: Progress Report. By this second milestone, you should have some initial results to share; for

example, you may have implemented and evaluated the performance of existing algorithms on your dataset and task of interest, or you may have conducted an initial study with simulated data to better understand the properties of certain methods, or you were able to prove some preliminary result about some question of interest, etc.

Your progress report deliverable is a write-up of no more than 2 pages (not including references) describing what you have accomplished so far and, briefly, what you intend to do in the remainder of the term. You should be able to reuse at least part of the text of this milestone in your final report.

III: In class presentation. You will present your work in class using your choice of (a combination of) the blackboard or a computer presentation. Paper handouts are optional. The length of the presentation will vary according to course enrollment, but each person should expect to speak for about 10-20 minutes. Since the talk will be short, you should consider practicing (and timing) your talk in advance. You can practice part of it with me in office hours if you want. Expect around 3-5 minutes of questions from myself and your fellow students. Unlike other times in the class, attendance is mandatory during the presentations, and I will be taking attendance during them. Once I set the time limits they will be strict. I will indicate when you go over time. Going over time will result in penalties.

You will be graded on your presentation skills, e.g. voice volume, board usage, pacing of material, choice of material, etc. Minor technical problems will not be penalized, but major technical problems will be penalized. If you use a computer, make sure to have a backup plan in case of a technical problem to avoid such a penalty. If you want to give a short version of your presentation in office hours before the actual presentation, and then have me give feedback that might be a good idea.

IV: Final Report. Your final project report (not including acknowledgements and references) should be around 5-8 pages in length (using at most 12 point font and maximum 1 inch margins) and should follow a typical scientific style (with abstract, introduction, etc.). The write-up should clearly define your problem or question of interest, review relevant past work, and introduce and detail your approach. A comprehensive empirical evaluation could follow, or some proofs of some results, along with an interpretation of your results. Any elucidation of the theoretical properties of an empirical method under consideration is also welcome.

If this work was done in collaboration with someone outside of the class (e.g., a professor), please describe their contributions in an acknowledgements section.

The final report PDF file should be emailed. No hardcopy is needed.

**Student Conduct:** 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 <http://equity.usc.edu/> or to the Department of Public Safety <http://capsnet.usc.edu/department/departement-public-safety/online-forms/contact-us>. This is important for the safety whole USC community. Another member of 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 <http://www.usc.edu/student-affairs/cwm/> provides 24/7 confidential support, and the sexual assault resource center webpage [sarc@usc.edu](mailto:sarc@usc.edu) describes reporting options and other resources.

**Accessibility Services:** If you are registered with accessibility services, I would be happy to discuss this at the beginning of the course. Any student requesting accommodations based on a disability is required to register with Accessibility Services and Programs (OSAS) each semester. A letter of verification for approved accommodations can be obtained from OSAS. Please be sure the letter is delivered to me as early in the semester as possible. OSAS is located in 301 STU and is open 8:30am-5:00pm, Monday through Friday.

<https://osas.usc.edu>

213-740-0776 (phone)

213-740-6948 (TDD only)

213-740-8216 (fax)

OSASFrontDesk@usc.edu

**Other Resources:** [An introduction to mathematical arguments](#)

### **Homework Policy:**

- Late homework is not accepted.
- If you still want to turn in late homework, then the number of minutes late, divided by ten, will be deducted from the score. (The time estimate is not guaranteed to be accurate.)
- Do not submit homework via email.
- There will be 3 homework assignments, turned in on blackboard at **9AM on Fridays**.
- A random subset of the homework problems will be graded each week. However, it is strongly recommended that you try to complete the entire homework assignment.
- You may use whatever resources you want to do the homework, including computers, textbooks, friends, the TA, etc. However, I would discourage any over-reliance on search technology such as Google, since its overuse could degrade your learning experience. By the end of the semester, you should be able to do the entire homework on your own, without any external help.
- All homework assignments must be **written by you**, i.e. you cannot copy someone else's solution verbatim. However, I would very much encourage you to form study groups and do the homework together in small groups. Homework is the most important part of a graduate mathematics course, and I encourage you to take it very seriously.
- Homework solutions will be posted on Friday after the homework is turned in.

### **Grading Policy:**

- The final course grade is weighted in the following way. Homework (45%), class participation (5%), project abstract (3%), project progress report (7%), final presentation (20%), final project report (30%).

- Class participation is not the same as attendance. Especially for graduate students, you have many things to do, and life happens, so I will never explicitly take attendance, but I will notice if someone is frequently absent. Things that increase your class participation grade include: asking good questions, paying attention in class, showing up on time or early to class, etc. Things that decrease your class participation grade include: excessive talking or disruptions during class, frequent absences, excessive texting/smartphone usage in class, frequent tardiness, etc.

**Tentative Schedule:** (This schedule may change slightly during the course.)

Week	Monday	Tuesday	Wednesday	Thursday	Friday
1	Aug 23: Introduction	Aug 24	Aug 25: Introduction, Vertex Cover, SVD, PCA	Aug 26	Aug 27: k-means, power method
2	Aug 30: Learning linear threshold functions	Aug 31	Sep 1: Perceptron Algorithm	Sep 2	Sep 3: Perceptron Algorithm
3	Sep 6: No class	Sep 7	Sep 8: Embeddings and the “kernel trick”	Sep 9	Sep 10: Homework 1 due. Mercer’s Theorem
4	Sep 13: Concentration of Measure, Hoeffding’s Inequality	Sep 14	Sep 15: Gaussian Concentration, Johnson-Lindenstrauss	Sep 16	Sep 17: Bennett’s Inequality
5	Sep 20: Empirical Risk Minimization	Sep 21	Sep 22: Gaussian Processes	Sep 23	Sep 24: Abstract due. Gaussian Processes
6	Sep 27: Sub-Gaussian Processes	Sep 28	Sep 29: General Empirical Processes	Sep 30	Oct 1: General Empirical Processes
7	Oct 4: PAC Learning	Oct 5	Oct 6: PAC Learning	Oct 7	Oct 8: Homework 2 due. Vapnik-Chernovenkis (VC) Theory
8	Oct 11: VC Theory	Oct 12	Oct 13: VC Theory	Oct 14	Oct 15: No class
9	Oct 18: VC Theory	Oct 19	Oct 20: Sparse recovery	Oct 21	Oct 22: high-dimensional convex geometry
10	Oct 25: Compressed sensing and LASSO	Oct 26	Oct 27: Low rank matrix recovery	Oct 28	Oct 29: Progress report due. Low rank matrix recovery
11	Nov 1: Boosting	Nov 2	Nov 3: Boosting	Nov 4	Nov 5: Homework 3 due. Deep Learning
12	Nov 8: Deep Learning	Nov 9	Nov 10: Deep Learning	Nov 11	Nov 12: Deep Learning
13	Nov 15: Review of course	Nov 16	Nov 17: Leeway	Nov 18	Nov 19: Leeway
14	Nov 22: Leeway	Nov 23	Nov 24: No class	Nov 25: No class	Nov 26: No class
15	Nov 29: Student Presentations	Nov 30	Dec 1: Student Presentations	Dec 2	Dec 3: Student Presentations

### Advice on succeeding in a math class:

- Review the relevant course material **before** you come to lecture. Consider reviewing course material a week or two before the semester starts.
- When reading mathematics, use a pencil and paper to sketch the calculations that are performed by the author.
- Come to class with questions, so you can get more out of the lecture. Also, finish your homework at least **two days** before it is due, to alleviate deadline stress.
- Write a rough draft and a separate final draft for your homework. This procedure will help you catch mistakes. Also, I would very much recommend [typesetting](#) your homework. Learning LaTeX is a very important skill to have for doing mathematics. [Here](#) is a template .tex file if you want to get started typesetting.
- If you are having difficulty with the material or a particular homework problem, review Polya's [Problem Solving Strategies](#), and come to office hours.