USC Viterbi School of Engineering

DSci 599: Machine Learning for a Secure Internet

Units: 4.0

Spring 2025 — Tuesday/Thursday — 8:00am–9:50am

Instructor: John Heidemann

Contact info: johnh@isi.edu

Location: GFS 207

Syllabus Changes

2025-01-09 none yet.

2025-01-13: added [Shen24a] about LLM jailbreaking.

2025-01-15: added office hours.

2025-02-14: brought [Fukuda17a], [Jacobs22a], and [Wright08a] each forward one week.

Catalog Course Description

Machine Learning for a Secure Internet: introduction to the design principles, layering, protocol design/analysis, networked applications, Internet structure/architecture, protocols for transport/congestion control, and network security. Application of ML algorithms for networking data and packet traces.

Recommended preparation: Familiarity with Python

Course Description

The goal of the course is to introduce students to state-of-the-art research on reasoning with network and cyber security data. The class projects will play a central role in the course to provide hands-on experience with applying ML algorithms to real-world networking data.

This course will teach data science students to collect, clean, and develop ML models for networking and cyber security data. The ML algorithms will include statistical learning, classification, clustering, link prediction, anomaly detection, Bayesian models, and similar algorithms. The course will enable students to:

- Develop understanding of ML methods for experimentation and correlational research for networking and cyber security
- Determine the required statistical analyses for modeling and evaluation of networked systems

The course topics will be particularly relevant to students interested in networked systems such as IoT and cellular networks and cyber security.

Learning Objectives

After completing this course, students will be able to:

Design and Execute Networking and Cyber Security Experiments

- describe how setup a networking experiment and collect data
- describe how setup a cyber security experiment and collect data
- describe how clean and process the network and log data
- choose the appropriate research design for their research questions, while considering both validity and rigor
- plan and evaluate choices in determining the specific design of an experiment

Application of ML Algorithms on Networking and Cyber Security Experiments

select the appropriate ML algorithms for the experiment

run the statistical test and understand the output

write up and interpret the results of the statistical tests

conduct analysis to determine the likelihood of finding a significant results

Technological Proficiency and Hardware/Software Required

All students will need introductory familiarity with Python programming.

Required Readings and Supplementary Materials

The primary required readings will be papers provided on the class website.

Two textbooks provide supplementary background on networking and ML:

Title:	Computer Networks: A Systems Approach
Authors:	Larry Peterson and Bruce Davie
Copyright:	Elsevier, 2012
Source:	https://github.com/SystemsApproach/book
License:	CC BY 4.0

and

Title:	Machine Learning: A Probabilistic Perspective
Authors:	Kevin P. Murphy
Copyright:	MIT Press, 2012

Grading Breakdown

midterm	25%
final	30%
homework	15%
projects	25%
class participation	5%

Course Evaluations

Course evaluation occurs at the end of the semester university-wide. It is an important review of students' experience in the class. The process and intent of the end-of-semester evaluation should be provided. In addition, a mid-semester evaluation is recommended practice for early course correction. You may choose to contact CET for support in creating a mid-semester evaluation.

Course Schedule

Class meets Tuesday and Thursday, 8am to 9:50am, beginning Tuesday, 2025-01-14 and ending Thursday, 2025-05-01.

We will have a **midterm exam** at 8am Tuesday 2025-02-25, and a **final exam** on Wednesday 2025-05-14, from 8:00 to 10:00am.

All students are expected to confirm they can make both the midterm and final exams—we do not offer alternative dates.

I expect topics will shift some over the semester, and I will add some papers. Any changes will be announced on the Moodle.

Office Hours: (added 2025-01-15) Office hours are Tuesdays, 10am to noon in GCS LL2. The professors campus office this semester is GCS 302F, but most students don't have direct access–e-mail meto schedule a meeting time and I can come get you.

Week 1: Introduction, Background, and Reading Papers

(Jan. 14 and Jan. 16)

Introduction to networking and ML, and how to read papers.

Tips for reading papers: [Hanson99a]

P1. **[Hanson99a]** Michael J. Hanson. Efficient reading of papers in science. Brochure of unknown origin, revised 1999 by Dylan J. McNamee, 1989.

Another viewpoint of paper reading [Jamin03a]

P2. **[Jamin03a]** Sugih Jamin. Paper reading and writing check lists. web page http://irl.eecs.umich.edu/jamin/courses/eecs589/papers/checklist.html, November 2003.

NABC-format paper presentations.

No paper, but we will review and discuss: General networking and machine learning.

Week 2: Background, Presenting Work, and Ethics

(Jan. 21 and Jan. 23)

Introduction to evaluating papers, ethics, and more background about networking and ML.

Finding and judging new ideas: [Heilmeier92a]

P3. **[Heilmeier92a]** George H. Heilmeier. Some reflections on innovation and invention. *The Bridge*, 22:12–16, Winter 1992.

Ethics: [Dittrich11a]

P4. **[Dittrich11a]** David Dittrich and Erin Kenneally (editors). The Menlo report: Ethical principles guiding information and communication technology research. Technical report, United States Department of Homeland Security, September 2011.

Week 3: Spam Classification and Bayesian Filtering

(Jan. 28 and Jan. 30)

Spam and Bayesian filtering.

A Plan for Spam (an informal proposal): [Graham02b]

P5. **[Graham02b]** Paul Graham. A plan for spam. Blog Post https://www.paulgraham.com/spam.html, August 2002.

[Sahami98a]

P6. [**Sahami98a**] Mehran Sahami, Susan Dumais, David Heckerman, and Eric Horvitz. A Bayesian approach to filtering junk e-mail. In *Proceedings of the AAAI Workshop on Learning for Text Categorization*, pages 55–62, Madison, Wisconsin, USA, July 1998. Assocation for hte Advancaement of Artificial Intelligence.

[Pantel98a]

P7. [**Pantel98a**] Patrick Pantel and Dekang Lin. Spamcop: A spam classification and organization program. In *Proceedings of the AAAI Workshop on Learning for Text Categorization*, pages 95–98, Madison, Wisconsin, USA, July 1998. Assocation for hte Advancaement of Artificial Intelligence.

Better Bayesian: [Graham03a]

P8. [**Graham03a**] Paul Graham. Better bayesian filtering. Blog Post https://www.paulgraham.com/better.html, January 2003.

Week 4: an End-to-End View of Spam

(Feb. 4 and Feb. 6)

A more complete view of the spam world and the many roles ML plays in analyzing it.

[Levchenko11a]

P9. [**Levchenko11a**] Kirill Levchenko, Andreas Pitsillidis, Neha Chachra, Brandon Enright, Márk Félegyházi, Chris Grier, Tristan Halvorson, Chris Kanich, Christian Kreibich, He Liu, Damon McCoy, Nicholas Weaver, Vern Paxson, Geoffrey M. Voelker, and Stefan Savage. Click trajectories: End-to-end analysis of the spam value chain. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 431–446, Oakland, CA, USA, May 2011. IEEE.

Week 5: Modeling, Bayesian Inference, and Outage Detection

(Feb. 11 and Feb. 13)

Modeling and framing detection via Bayesian inference, with network outages as an application.

[Quan13c]

P10. [**Quan13c**] Lin Quan, John Heidemann, and Yuri Pradkin. Trinocular: Understanding Internet reliability through adaptive probing. In *Proceedings of the ACM SIGCOMM Conference*, pages 255–266, Hong Kong, China, August 2013. ACM.

Week 6: Classification and Feature Selection and Model Drift

(Feb. 18 and Feb. 20)

ML-based classification and feature selection, applied to interpreting DNS data.

[Fukuda15a]

P11. [**Fukuda15a**] Kensuke Fukuda and John Heidemann. Detecting malicious activity with DNS backscatter. In *Proceedings of the ACM Internet Measurement Conference*, pages 197–210, Tokyo, Japan, October 2015. ACM.

Model drift in ML-based classification. (*Change 2025-02-14:* brought forward one week to here.)

[Fukuda17a]

P12. [**Fukuda17a**] Kensuke Fukuda, John Heidemann, and Abdul Qadeer. Detecting malicious activity with DNS backscatter over time. *ACM/IEEE Transactions on Networking*, 25(5):3203–3218, August 2017.

Week 7: Decision Trees and Learning from ML

(Feb. 25 and Feb. 27)

We will have a **midterm exam** at 8am Tuesday 2025-02-25. It will not be the full class period, and we will have lecture in the second half of class.

~~Model drift in ML-based classification.~~

The challenge of applying ML to sensitive topics, and using ML to create decision trees. (*Change 2025-02-14:* brought forward one week to here.)

[Jacobs22a]

P13. [**Jacobs22a**] Arthur S. Jacobs, Roman Beltiukov, Walter Willinger, Ronaldo A. Ferreira, Arpit Gupta, and Lisandro Z. Granville. AI/ML for network security: The emperor has no clothes. In *Proceedings of the ACM Conference on Computer and Communications Security*, pages 1537–1551, Los Angeles, CA, USA, November 2022. ACM.

Week 8: Seeing Trhough Encryption (start)

(Mar. 4 and Mar. 6)

~~The challenge of applying ML to sensitive topics, and using ML to create decision trees.~~

Use of ML to infer the contents of encrypted data (without directly attacking the encryption). (*Change 2025-02-14:* brought forward one week to here.)

[Wright08a]

P14. [**Wright08a**] Charles V. Wright, Lucas Ballard, Scott E. Coull, Fabian Monrose, and Gerald M. Masson. Spot me if you can: Uncovering spoken phrases in encrypted VoIP conversations. In *Proceedings of the IEEE Symposium on Security and Privacy*, Oakland, CA, USA, May 2008. IEEE.

Week 9: Seeing Through Encryption

(Mar. 11 and Mar. 13)

Use of ML to infer the contents of encrypted data (without directly attacking the encryption). (*Change 2025-02-14:* some of this content was moved to the week of Mar. 4.)

~~[Wright08a]~~

[Chen10a]

P15. [**Chen10a**] Shuo Chen, Rui Wang, XiaoFeng Wang, and Kehuan Zhang. Side-channel leaks in web applications: A reality today, a challenge tomorrow. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 191–206, Oakland, CA, USA, May 2010. IEEE.

Spring Break

Spring break: Mar. 18, Mar. 20.

Week 10: Hypothesis Testing and Parametric Modeling

(Mar. 25 and Mar. 27)

Confusion matrices, modeling traffic, and hypothesis testing, applied to detecting Distributed Denial-of-Service traffic.

[Thatte10a]

P16. **[Thatte10a]** Gautam Thatte, Urbashi Mitra, and John Heidemann. Parametric methods for anomaly detection in aggregate traffic. *ACM/IEEE Transactions on Networking*, 19(2):512–525, August 2010. (Appeared in print April 2011).

Week 11: Hierarchical Clustering in Machine Learning

(Apr. 1 and Apr. 3)

Unsupervised learning and hierarchical clustering, applied to evaluation of IP address usage.

[Cai10a]

P17. **[Cai10a]** Xue Cai and John Heidemann. Understanding block-level address usage in the visible Internet. In *Proceedings of the ACM SIGCOMM Conference*, pages 99–110, New Delhi, India, August 2010. ACM.

Week 12: Security Thinking Applied to Jailbreaking LLMs

(Apr. 8 and Apr. 10)

LLMs, protecting them from bad behavior, and attacking those protections ("jailbreaking").

[Shen24a]

P18. **[Shen24a]** Xinyue Shen, Zeyuan Chen, Michael Backes, Yun Shen, and Yang Zhang. "do anything now": Characterizing and evaluating in-the-wild jailbreak prompts on large language models. In *Proceedings of the ACM Conference on Computer and Communications Security*, pages 1671–1685. ACM, December 2024.

Week 13: Optimization and Congestion Control

(Apr. 15 and Apr. 17)

ML-based optimization, applied to congestion control.

[Winstein13a]

P19. **[Winstein13a]** Keith Winstein and Hari Balakrishnan. TCP ex machina: Computer-generated congestion control. In *Proceedings of the ACM SIGCOMM Conference*, pages 123–134, Hong Kong, China, August 2013. ACM.

Week 14: Student Project Presentations

(Apr. 22 and Apr. 24)

Student project presentations.

Week 15: Student Project Presentations

(Apr. 29 and May. 1)

Student project presentations.

If all students agree, potentially we can do the final exam in the last class period.

Finals Week

The final exam is Wednesday 2025-05-14, from 8:00 to 10:00am (sorry, it's the University's choice of start time).

Academic Integrity

USC's Academic Integrity Policy

The University of Southern California is foremost a learning community committed to fostering successful scholars and researchers dedicated to the pursuit of knowledge and the transmission of ideas. Academic misconduct—which includes any act of dishonesty in the production or submission of academic work (either in draft or final form)—is in contrast to the university's mission to educate students through a broad array of academic, professional, and extracurricular programs.

This course will follow the expectations for academic integrity as stated in the USC Student Handbook (https://policy.usc.edu/studenthandbook/). All students are expected to submit assignments that are their own original work and prepared specifically for this course and section in this academic term. You may not submit work written by others or "recycle" work prepared for other courses without obtaining written permission from the instructor(s). Students suspected of engaging in academic misconduct will be reported to the Office of Academic Integrity.

Other violations of academic misconduct include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

Academic dishonesty has a far-reaching impact and is considered a serious offense against the university. Violations will result in a grade penalty, such as a failing grade on the assignment or in the course, and disciplinary action from the university itself, such as suspension or even expulsion.

For more information about academic integrity see the student handbook (https://policy.usc.edu/studenthandbook/) or the Office of Academic Integrity's website (https://academicintegrity.usc.edu/), and university policies on Research and Scholarship Misconduct (https://policy.usc.edu/research-and-scholarship-mi

Please ask the instructor [and/or TA(s)] if you are unsure about what constitutes unauthorized assistance on an exam or assignment, or what information requires citation and/or attribution.

Use of Generative AI in this Course

Generative AI is not permitted: Since creating, analytical, and critical thinking skills are part of the learning outcomes of this course, all assignments should be prepared by the student working individually or in groups as described on each assignment. Students may not have another person or entity complete any portion of the assignment. Developing strong competencies in these areas will prepare you for a competitive workplace. Therefore, using AI-generated tools is prohibited in this course, will be identified as plagiarism, and will be reported to the Office of Academic Integrity.

Class Recordings and Course Content Distribution

You may not record this class without the express permission of the instructor and all other students in the class. Distribution of any notes, recordings, exams, or other materials from a university class or lectures—other than for individual or class group study—is prohibited without the express permission of the instructor; violations will be considered an intentional act to facilitate or enable academic dishonesty and reported to the university

Course Content Distribution and Synchronous Session Recordings Policies

USC has policies that prohibit recording and distribution of any synchronous and asynchronous course content outside of the learning environment.

Recording a university class without the express permission of the instructor and announcement to the class, or unless conducted pursuant to an Office of Student Accessibility Services (OSAS) accommodation. Recording can inhibit free discussion in the future, and thus infringe on the academic freedom of other students as well as the instructor. (Living our Unifying Values: The USC Student Handbook, page 13, https://policy.usc.edu/studenthandbook/).

Distribution or use of notes, recordings, exams, or other intellectual property, based on university classes or lectures without the express permission of the instructor for purposes other than individual or group study. This includes but is not limited to providing materials for distribution by services publishing course materials. This restriction on unauthorized use also applies to all information, which had been distributed to students or in any way had been displayed for use in relation to the class, whether obtained in class, via email, on the internet, or via any other media. Distributing course material without the instructor's permission will be presumed to be an intentional act to facilitate or enable academic dishonestly and is strictly prohibited. (Living our Unifying Values: The USC Student Handbook, page 13, https://policy.usc.edu/studenthandbook/)

Statement on University Academic and Support Systems

Students and Disability Accommodations

USC welcomes students with disabilities into all of the University's educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Student Financial Aid and Satisfactory Academic Progress

To be eligible for certain kinds of financial aid, students are required to maintain Satisfactory Academic Progress (SAP) toward their degree objectives. Visit the Financial Aid Office webpage for undergraduate- and graduate-level SAP eligibility requirements and the appeals process.

Support Systems

Counseling and Mental Health: (213) 740–9355: 24/7 on call.

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

988 Suicide and Crisis Lifeline: 988 for both calls and text messages: 24/7 on call.

The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline consists of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273–8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

Relationship and Sexual Violence Prevention Services (RSVP): (213) 740-9355(WELL): 24/7 on call

Free and confidential therapy services, workshops, and training for situations related to gender- and power- based harm (including sexual assault, intimate partner violence, and stalking).

Office for Equity, Equal Opportunity, and Title IX (EEO-TIX): (213) 740-5086

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

Reporting Incidents of Bias or Harassment: (213) 740-2500

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

The Office of Student Accessibility Services (OSAS): (213) 740-0776

OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

USC Campus Support and Intervention: (213) 740-0411

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity, Equity and Inclusion: (213) 740-2101

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency UPC: (213) 740-4321, HSC: (323) 442-1000: 24/7 on call

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety: UPC: (213) 740-6000, HSC: (323) 442-1200: 24/7 on call

Non-emergency assistance or information.

Office of the Ombuds: (213) 821-9556 (UPC) / (323-442-0382 (HSC)

A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

Occupational Therapy Faculty Practice: (323) 442-2850 or otfp@med.usc.edu

Confidential Lifestyle Redesign services for USC students to support health-promoting habits and routines that enhance quality of life and academic performance.