9.520/6.860: Statistical Learning Theory and Applications

- Class: **Tue, Thu 11:00 am - 12:30 pm**, 46-3002 Singleton Auditorium
- Office Hours: **Wednesday 3:00pm – 4:00pm (in person at 46-5193)**
- Web: https://cbmm.mit.edu/9-520
- Contact: 9.520@mit.edu
- 9.520/6.860 will use Canvas: https://canvas.mit.edu/courses/16755
- Also check Canvas announcements for updates
- **This year's course will be in-person until MIT policy changes.**
- **Please fill out this registration form at https://forms.gle/sz8eX9pnjopeBJHp6**

Material

Slides— will be posted (for most lectures) on the website and Canvas

Videos— Recordings of lectures will be made available on Canvas

Notes—

L. Rosasco and T. Poggio, **Machine Learning: a Regularization Approach, MIT-9.520 Lectures Notes**, Manuscript, **(will be provided)**

For feedback on book (typos, errors, ...)
https://goo.gl/forms/pQcewnsAV3lCNoyr1

Faces

- Instructors:

Faces

- Instructors:
  - Lorenzo Rosasco

Faces

- Instructors:
  - Lorenzo Rosasco
  - Tomaso Poggio

Faces

- Instructors:
  - Lorenzo Rosasco
  - Tomaso Poggio
  - Akshay Rangamani (head TA 1)

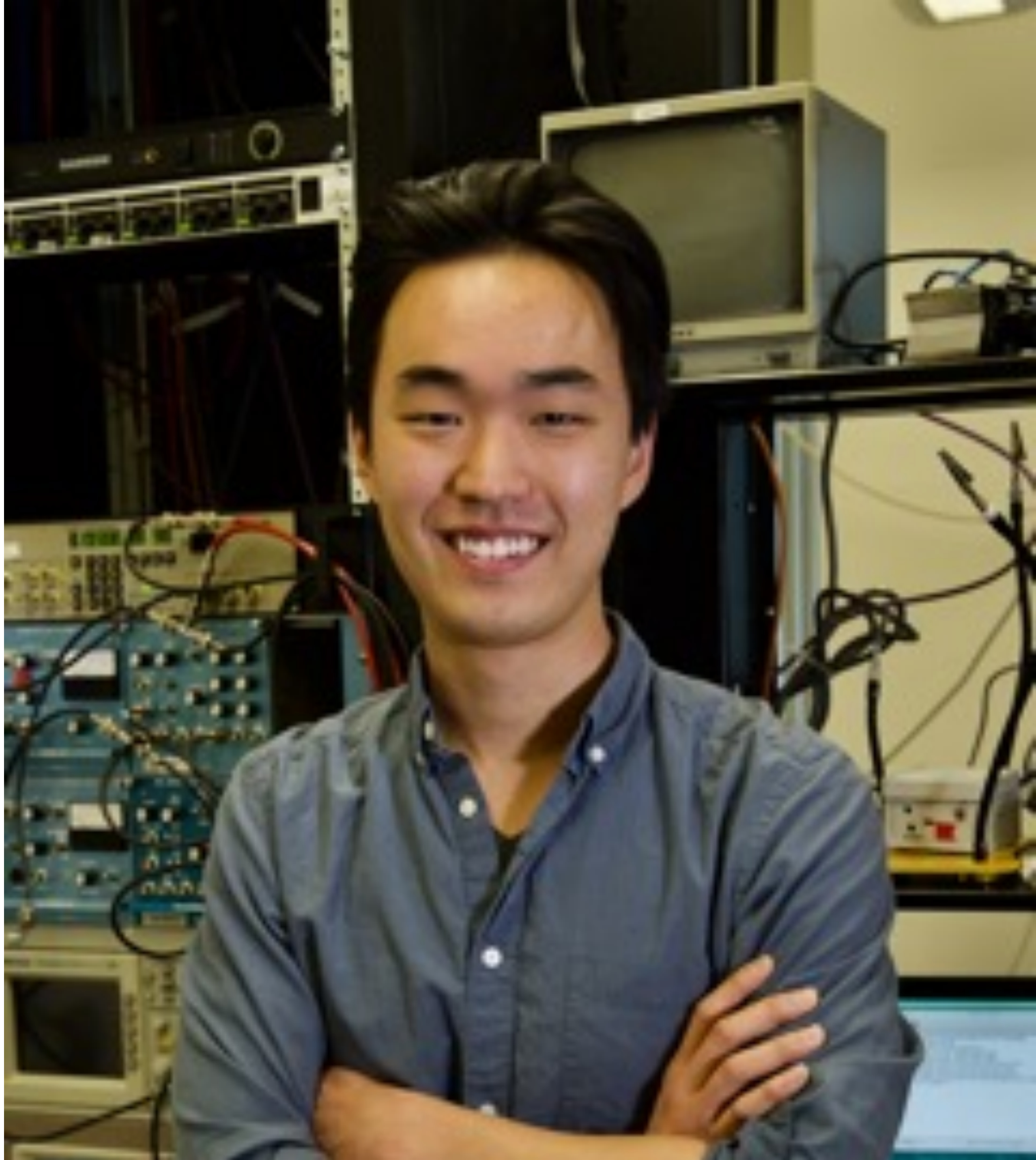
Faces

- Instructors:
  - Lorenzo Rosasco
  - Tomaso Poggio
  - Akshay Rangamani (also head TA 1)
  - Tomer Galanti (also head TA 2)

Faces

- Instructors:
  - Lorenzo Rosasco
  - Tomaso Poggio
  - Akshay Rangamani (also head TA 1?)
  - Tomer Galanti (also head TA 2?)

- TAs:
  - Michael Lee



Other Machine Learning courses at MIT

- Introduction to Machine Learning (6.036)
  - “Introduces principles, algorithms, and applications of machine learning … methods such as support vector machines, hidden Markov models, and neural networks.”

- Machine Learning (6.867)
  - “Principles, techniques, and algorithms in machine learning … linear/additive models, active learning, boosting, support vector machines, non-parametric Bayesian methods, hidden Markov models, Bayesian networks”

- This course (9.520/6.860)
  - “Among different approaches in modern machine learning, the course focuses on a regularization perspective … It will introduce an emerging theoretical framework addressing three key puzzles in deep learning: approximation theory -- which functions can be represented more efficiently by deep networks than shallow networks -- optimization theory -- why can stochastic gradient descent easily find global minima -- and machine learning -- whether classical learning theory can explain generalization in deep networks.”

Syllabus at a glance

Class	Date	Title	Instructor(s)
Class 01	Thu Sep 08	Course Outline. Statistical Machine Learning	LR
Class 02	Tue Sep 13	Empirical Risk Minimization and Regularization for Linear Models	LR
Class 03	Thu Sep 15	Kernels and Feature Maps	LR
Class 04	Tue Sep 20	Optimization: GD and SGD. Regularization and implicit regularization	TG
Class 05	Thu Sep 22	Error Decomposition and Approximation Error	AR
Class 06	Tue Sep 27	Estimation Error and Generalization Gap	AR
Class 07	Thu Sep 29	Stability of Ridge and Ridgeless Regression	AR
Class 08	Tue Oct 04	Introduction to Deep Networks	AR
Class 09	Thu Oct 06	Deep Learning Theory: Approximation	TP
Monday 10th October - Indigenous People's Day, Tuesday 11th October - Student Holiday			
Class 10	Thu Oct 13	Deep Learning: Optimization and Dynamics	TP
Class 11	Tue Oct 18	Deep Learning: Bias towards Low Rank	TG
Class 12	Thu Oct 20	Deep Learning: Neural Collapse	AR + TG
Class 13	Tue Oct 25	Deep Learning: Generalization	TP
Class 14	Thu Oct 27	Group Invariance and Equivariance in Vision and Learning	Fabio Anselmi
Class 15	Tue Nov 01	Brain and Neural Networks - Identification Problems	Brian Cheung + Yena Han
Class 16	Thu Nov 03	Neural Networks and the Ventral Stream	Thomas Serre + Gabriel Kreiman
Class 17	Tue Nov 08	Loose Ends	Staff
Class 18	Tue Nov 15	Sample and computational complexity of deep networks	TBD
Class 19	Thu Nov 17	Overparameterized in Learning	TBD
Class 20	Tue Nov 22	Transformers	Brian Cheung
Thursday 24th November - Thanksgiving			
Class 21	Tue Nov 29	Neural Assemblies	Christos Papadimitriou +Santosh Vempala
Class 22	Thu Dec 01	Adversarial examples	TBD
Class 23	Tue Dec 06	The loss landscape of overparametrized deep nets	Yaim Cooper
Class 24	Thu Dec 08	Transformers	TBD
Class 25	Tue Dec 13	Transformers	Tomer Ullman

Grading policies

Problem sets (0.45)

- 3 problem sets (0.15 each)
  - 3 - 4 questions (exercises and/or MATLAB)
  - Due in 14 days (out on a class day, and due before class two weeks later)
- Late policy on next slide
- typeset in LaTeX (template will be provided)
- Online submission on Canvas by due date

Project (0.45)

- See later

Class Participation (0.10)

- *Attending class lectures is required!*
- Discussions during class - ask questions!
- Discussions on Canvas

Problem sets

- Problem sets (0.45)
  - 3 problem sets (0.15 each)
    - 3 - 4 questions (demonstrations/exercises)
    - 10 days due!
  - typeset in LaTeX (template provided)
  - *online submission on Canvas by due date*
- ***Late policy***
  - All students have 6 free late days (to be used on psets and project proposal)
  - You may use them as you see fit
  - Beyond this, we will not accept assignments
- **Dates (due times are 10:59 am). Submission online (on Canvas).**
  - **Problem Set 1, out: Tue. Sept. 20, due: Tue. Oct. 04**
  - **Problem Set 2, out: Thu. Oct. 13, due: Thu. Oct. 27**
  - **Problem Set 3, out: Tue. Nov. 01, due: Tue. Nov. 15**
- **Collaboration policy: You may discuss with others but need to work out your own solution.**

Projects

Theory

Algorithms

Review

Application

- This is not a data science course, so we will not consider data preparation as contributing to the grade.

Final Evaluation: project report (5 pages for individuals, 8 pages for teams, NeurIPS style)

Dates

- Abstract and title: Oct. 20
- Feedback and approval: Oct. 27
- Final Report due: Dec.13