

# CS189 - Introduction to Machine Learning

CS 189 / 289A

Introduction to Machine Learning

**Instructors:** **Jitendra Malik** (<https://people.eecs.berkeley.edu/~malik/>) and **Benjamin Recht** (<http://eecs.berkeley.edu/~brecht>)

**Time:** TuTh 12:30-2:00 PM

**Location:** 245 Li Ka Shing

**GSIs:** Anurag Ajay, Brian Chu, Daylen Yang, Garrett Thomas, Giulio Zhou, Nick Boyd, Samaneh Azadi, Smitha Milli, Stephen Tu, Vaishaal Shankar, Yiqun Chen

**Piazza Site:** <https://piazza.com/berkeley/fall2016/cs189> (<https://piazza.com/berkeley/fall2016/cs189>)



This course provides an introduction to the theoretical foundations, algorithms, and applications of contemporary machine learning. We will study the methods to find patterns in data and how to use them to make predictions. We will describe how machine learning rests on connections between probabilistic models, optimization problems, and nonlinear programming algorithms. A tentative list of topics includes

- **Classification**
  - Risk minimization and ML Abstractions
  - Stochastic Gradient Descent
- **Decision Theory**
  - Multivariate Gaussians
  - Eigenvalues and eigenvectors
  - Anisotropic normal distributions
- **Regression**
  - motivations from curve fitting and statistics
  - practical algorithms, normal equations and gradient descent
- **Regularization**
  - The Bias Variance Tradeoff
  - Ridge Regression and Tikhonov Regularization
  - Feature selection and the LASSO
- **Neural Networks**
  - Nonlinear regression, feature maps and feature engineering
  - Neural Networks and Backpropagation
  - Motivations from neuroscience
- **Nonparametric methods**
  - Nearest Neighbors
  - High-dimensional data analysis
  - Kernel Methods

- Decision Trees
- Random Forests and Ensembling
- **Unsupervised learning**
  - k-means and vector quantization
  - principal components analysis
  - the singular value decomposition
  - spectral graph partitioning and graph clustering

### Recommended texts:

Both textbooks for this class are available free online. Hardcover and Kindle/eTextbook versions are also available.

- Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, *[An Introduction to Statistical Learning with Applications in R](http://www-bcf.usc.edu/~gareth/ISL/)* (<http://www-bcf.usc.edu/~gareth/ISL/>), Springer, New York, 2013. ISBN # 978-1-4614-7137-0. See Amazon for hardcover or eTextbook.
- Trevor Hastie, Robert Tibshirani, and Jerome Friedman, *[The Elements of Statistical Learning: Data Mining, Inference, and Prediction](http://statweb.stanford.edu/~tibs/ElemStatLearn/)* (<http://statweb.stanford.edu/~tibs/ElemStatLearn/>), second edition, Springer, 2008. See Amazon for hardcover or Kindle.
- **Jonathan Shewchuk's Lecture notes**  
<https://bcourses.berkeley.edu/courses/1454815/files/69495280/download>   
<https://bcourses.berkeley.edu/courses/1454815/files/69495280/download>   
<https://bcourses.berkeley.edu/courses/1454815/files/69495280/download>

### Prerequisites

- Math 53 (vector calculus)
- Math 54 (linear algebra)
- CS 70 (discrete math; probability)

### Grading

- 40% Homework: 6 assignments. Late policy: 5 slip days total.
- 20% Midterm: Tuesday, October 18, 12:30-2pm
- 40% Final Exam: Friday, December 16, 8-11am (Exam group 17)

### Course rules:

- Discussion of HW problems is encouraged.
- All homeworks, including programming, must be written individually.
- We will actively check for plagiarism.
- Typical penalty is a large NEGATIVE score, but we reserve right to give an instant F for even one violation, and will always give an F for two.
- You have 5 slip days total to use over the course of the semester on homeworks. These are meant to

be used in EMERGENCIES ONLY. If you use all your slip days halfway through the semester and then happen to have a real emergency, we can't help you. Ideally, you should end the semester without having used your slip days.











## Office Hours


















[https://calendar.google.com/calendar/embed?mode=WEEK&src=berkeley.edu\\_k02d2r1oh3q1kvnrtdvffb929o%40group.calendar.google.com&ctz=\(https://calendar.google.com/calendar/embed?mode=WEEK&src=berkeley.edu\\_k02d2r1oh3q1kvnrtdvffb929o%40group.calendar.google.com&ctz=America/L](https://calendar.google.com/calendar/embed?mode=WEEK&src=berkeley.edu_k02d2r1oh3q1kvnrtdvffb929o%40group.calendar.google.com&ctz=(https://calendar.google.com/calendar/embed?mode=WEEK&src=berkeley.edu_k02d2r1oh3q1kvnrtdvffb929o%40group.calendar.google.com&ctz=America/L)






## Discussion Sections

[https://calendar.google.com/calendar/embed?mode=WEEK&src=berkeley.edu\\_71p1e5i423m6j0csp6gakib07o%40group.calendar.google.com&ctz=\(https://calendar.google.com/calendar/embed?mode=WEEK&src=berkeley.edu\\_71p1e5i423m6j0csp6gakib07o%40group.calendar.google.com&ctz=America/L](https://calendar.google.com/calendar/embed?mode=WEEK&src=berkeley.edu_71p1e5i423m6j0csp6gakib07o%40group.calendar.google.com&ctz=(https://calendar.google.com/calendar/embed?mode=WEEK&src=berkeley.edu_71p1e5i423m6j0csp6gakib07o%40group.calendar.google.com&ctz=America/L)

## Lectures

	Date	Topic	Reading
1	8/25	Machine Learning Concepts and Data Hygiene	<a href="#">lecture1 (powerpoint)</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69522103/download">https://bcourses.berkeley.edu/courses/1454815/files/69522103/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69522103/download">https://bcourses.berkeley.edu/courses/1454815/files/69522103/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69522103/download">https://bcourses.berkeley.edu/courses/1454815/files/69522103/download</a>
2	8/30	Classification	Shewchuk Lectures 2 and 3, ESL Section 4.5, <a href="#">Ben's notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69532391/download">https://bcourses.berkeley.edu/courses/1454815/files/69532391/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69532391/download">https://bcourses.berkeley.edu/courses/1454815/files/69532391/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69532391/download">https://bcourses.berkeley.edu/courses/1454815/files/69532391/download</a>
3	9/1	Unconstrained optimization and gradient descent	<a href="#">SGD Notes.pdf</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69562506/download">https://bcourses.berkeley.edu/courses/1454815/files/69562506/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69562506/download">https://bcourses.berkeley.edu/courses/1454815/files/69562506/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69562506/download">https://bcourses.berkeley.edu/courses/1454815/files/69562506/download</a> . Shewchuck Lecture 5, <a href="#">Ben's Notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69574250/download">https://bcourses.berkeley.edu/courses/1454815/files/69574250/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69574250/download">https://bcourses.berkeley.edu/courses/1454815/files/69574250/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69574250/download">https://bcourses.berkeley.edu/courses/1454815/files/69574250/download</a>
4	9/6	stochastic gradient descent	The SGD notes and Shewchuk lecture 3, <a href="#">Ben's Notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69574249/download">https://bcourses.berkeley.edu/courses/1454815/files/69574249/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69574249/download">https://bcourses.berkeley.edu/courses/1454815/files/69574249/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69574249/download">https://bcourses.berkeley.edu/courses/1454815/files/69574249/download</a>
			Shewchuck Lecture 4 and 5, <a href="#">Ben's Notes</a>

5	9/8	RISK minimization and optimization abstractions	<a href="https://bcourses.berkeley.edu/courses/1454815/files/69587806/download">https://bcourses.berkeley.edu/courses/1454815/files/69587806/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69587806/download">https://bcourses.berkeley.edu/courses/1454815/files/69587806/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69587806/download">https://bcourses.berkeley.edu/courses/1454815/files/69587806/download</a>
6	9/13	Decision Theory	Shewchuk Lecture 6, <a href="#">Jitendra's slides</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69615737/download">https://bcourses.berkeley.edu/courses/1454815/files/69615737/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69615737/download">https://bcourses.berkeley.edu/courses/1454815/files/69615737/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69615737/download">https://bcourses.berkeley.edu/courses/1454815/files/69615737/download</a>
7	9/15	Multivariate Gaussians and Random Vectors	Shewchuk Lectures 8 and 9. <a href="#">Ben's Notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69664915/download">https://bcourses.berkeley.edu/courses/1454815/files/69664915/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69664915/download">https://bcourses.berkeley.edu/courses/1454815/files/69664915/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69664915/download">https://bcourses.berkeley.edu/courses/1454815/files/69664915/download</a>
8	9/20	Maximum Likelihood	Shewchuk Lecture 7, ISL Section 4.4 and 4.5. <a href="#">Ben's Notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69664917/download">https://bcourses.berkeley.edu/courses/1454815/files/69664917/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69664917/download">https://bcourses.berkeley.edu/courses/1454815/files/69664917/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69664917/download">https://bcourses.berkeley.edu/courses/1454815/files/69664917/download</a>
9	9/22	LDA and QDA	Shewchuck Lectures 8 and 9. <a href="#">Ben's Notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69699159/download">https://bcourses.berkeley.edu/courses/1454815/files/69699159/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69699159/download">https://bcourses.berkeley.edu/courses/1454815/files/69699159/download</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69699159/download">https://bcourses.berkeley.edu/courses/1454815/files/69699159/download</a>
10	9/27	Regression	Shewchuk Lecture 10. ISL Sections 4-4.3. notes on <a href="#">derivatives</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719835/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719835/download?wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719835/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719835/download?wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719835/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719835/download?wrap=1</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719836/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719836/download?wrap=1</a> , <a href="#">Linear regression notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719836/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719836/download?wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719836/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719836/download?wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719836/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719836/download?wrap=1</a> , <a href="#">Linear regression slides</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719837/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719837/download?wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719837/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719837/download?wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69719837/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69719837/download?wrap=1</a>
11	9/29	Logistic Regression and algorithms	Shewchuk Lecture 11. Sections 4.4.3, 7.1, 9.3.3; ESL, Section 4.4.1. Notes: <a href="#">Logistic-regression-notes.pdf</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69768666/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69768666/download?wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69768666/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69768666/download?wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69768666/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69768666/download?wrap=1</a> , <a href="#">Linear2+Logistic.pptx</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69768669/download?wrap=1">https://bcourses.berkeley.edu/courses/1454815/files/69768669/download?wrap=1</a>

			<a href="https://bcourses.berkeley.edu/courses/1454815/files/69768669/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69768669/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69768669/download?wrap=1">wrap=1</a>
12	10/4	bias-variance tradeoffs	Shewchuk Lecture 12, <a href="#">Ben's Notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69784883/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69784883/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69784883/download?wrap=1">wrap=1</a>
13	10/6	regularization	Shewchuk Lectures 13 and 14 (but none of the kernel stuff), <a href="#">Ben's Notes</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69784882/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69784882/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69784882/download?wrap=1">wrap=1</a>
14	10/11	cross validation	Voodoo cross validation: <a href="http://biorxiv.org/content/early/2016/06/19/059774">http://biorxiv.org/content/early/2016/06/19/059774</a> <a href="http://biorxiv.org/content/early/2016/06/19/059774">http://biorxiv.org/content/early/2016/06/19/059774</a>  How to win at machine learning without looking at the data: <a href="http://blog.mrtz.org/2015/03/09/competition.html">http://blog.mrtz.org/2015/03/09/competition.html</a> <a href="http://blog.mrtz.org/2015/03/09/competition.html">http://blog.mrtz.org/2015/03/09/competition.html</a>  <a href="http://www.nytimes.com/2015/06/04/technology/computer-scientists-are-astir-after-baidu-team-is-barred-from-ai-competition.html">http://www.nytimes.com/2015/06/04/technology/computer-scientists-are-astir-after-baidu-team-is-barred-from-ai-competition.html</a> <a href="http://www.nytimes.com/2015/06/04/technology/computer-scientists-are-astir-after-baidu-team-is-barred-from-ai-competition.html">http://www.nytimes.com/2015/06/04/technology/computer-scientists-are-astir-after-baidu-team-is-barred-from-ai-competition.html</a>  <b>Ben's notes</b> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69816843/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69816843/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69816843/download?wrap=1">wrap=1</a>
15	10/13	fairness in machine learning (not tested)	Approaching fairness: <a href="http://blog.mrtz.org/2016/09/06/approaching-fairness.html">http://blog.mrtz.org/2016/09/06/approaching-fairness.html</a>
16	10/20	neural networks 1	<b>Jitendra's notes</b> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69885071/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69885071/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69885071/download?wrap=1">wrap=1</a> , <a href="#">slides</a> <a href="https://bcourses.berkeley.edu/courses/1454815/files/69885073/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69885073/download?wrap=1">wrap=1</a>  <a href="https://bcourses.berkeley.edu/courses/1454815/files/69885073/download?wrap=1">wrap=1</a>

