Course Schedule (Evolving)

Lecture recordings from Echo360 can be accessed here (https://echo360.org/section/570de0f1-4d7a-4a68-a5ce-3cc36ccbce07/public). Handwritten lecture notes courtesy of Stephen Scarano can be accessed here (https://drive.google.com/drive/folders/1gTHtuIMz8073YINv-iuC72tJ2n9mTrIn).

Lecture	Day	Topic	Materials/Reading
1. 9/5	Tue	Course overview. Probability review. Linearity of expectation.	Slides. (./slides/lecture1/lecture1Annotated.pdf) Compressed slides. (./slides/lecture1/lecture1Compressed.pdf) Reading: MIT short videos and exercise probability (https://openlearninglibrary.mit.edu/courses/course-v1:OCW+6.042J+2T2019/course/) (go to Unit 4). Khan academy probability lessor (https://www.khanacademy.org/math/statistics-probability/probability-library) (a basic). Chapters 1-3 of Probability and Computing (https://www.cs.purdue.edu/homes/spa/courses/pg17/mu-book.pdf) with content excersises on basic probabiliy, expectation, variance, and concentration bounds.
		Randomized Methods, Sketching & Streaming	
2. 9/7	Thu	Estimating set size by counting duplicates. Markov's inequality. Random hashing for efficient lookup. Collision-free hashing.	Slides. (./slides/lecture2/lecture2Annotated.pdf) Compressed slides. (./slides/lecture2/lecture2Compressed.pdf) Reading: Chapters 1-3 of Probability ar Computing (https://www.cs.purdue.edu/homes/spa/courses/pg17/mu-book.pdf) v content and excersises on basic probabiliy, expectation, variance, and concentratio
3. 9/12	Tue	More random hashing: 2-level hashing. 2-universal and pairwise independent hashing.	Slides. (./slides/lecture3/lecture3Annotated.pdf) Compressed slides. (./slides/lecture3/lecture3Compressed.pdf) Reading: Chapter 2.2 (https://www.cs.cornell.edu/jeh/book.pdf) of *Foundations of Data Science* with conte Markov's inequality and Chebyshev's inequality. Exercises 2.1-2.6. Chapters 1-3 of and Computing (https://www.cs.purdue.edu/homes/spa/courses/pg17/mu-book.p content and excersises on basic probabiliy, expectation, variance, and concentratio Some notes (https://www.cs.princeton.edu/courses/archive/fall16/cos521/Lecture (Arora and Kothari at Princeton) proving that the ax+b mod p hash function describ in 2-universal.
4. 9/14	Thu	Hashing for load balancing. Chebyshev's inequality. The union bound.	Slides. (./slides/lecture4/lecture4Annotated.pdf) Compressed slides. (./slides/lecture4/lecture4Compressed.pdf) Reading: Chapter 2.2 (https://www.cs.cornell.edu/jeh/book.pdf) of *Foundations of Data Science* with conte Markov's inequality and Chebyshev's inequality. Exercises 2.1-2.6. Chapters 1-3 of and Computing (https://www.cs.purdue.edu/homes/spa/courses/pg17/mu-book.p content and excersises on basic probabiliy, expectation, variance, and concentratio
5. 9/19	Tue	Exponential concentration bounds and the central limit theorem.	Slides. (./slides/lecture5/lecture5Annotated.pdf) Compressed slides. (./slides/lecture5/lecture5Compressed.pdf) Reading: Chapter 4 of Probability and ((https://www.cs.purdue.edu/homes/spa/courses/pg17/mu-book.pdf) on exponent concentration bounds. Some notes (http://math.mit.edu/~goemans/18310S15/che notes.pdf) (Goemans at MIT) showing how to prove exponential tail bounds using t generating function + Markov's inequality approach discussed in class.
6. 9/21	Thu	Finish up exponential concentration bounds. Bloom Filters.	Slides. (./slides/lecture6/lecture6Annotated.pdf) Compressed slides. (./slides/lecture6/lecture6Compressed.pdf) Reading: Chapter 4 of Probability and ((https://www.cs.purdue.edu/homes/spa/courses/pg17/mu-book.pdf) on exponent concentration bounds. Chapter 4 (http://infolab.stanford.edu/~ullman/mmds/ch4. *Mining of Massive Datasets*, with content on bloom filters. See here (https://cstheory.stackexchange.com/questions/6596/a-probabilistic-set-with-no positives/14455#14455) for some explaination of why a version of a Bloom filter w negatives cannot be achieved without using a lot of space. See Wikipedia

			(https://en.wikipedia.org/wiki/Bloom_filter#Bloomier_filters) for a discussion of the bloom filter variants, including counting Bloom filters, and Bloom filters with deleti Wikipedia again (https://en.wikipedia.org/wiki/Cuckoo_hashing) and these notes (https://web.stanford.edu/class/archive/cs/cs166/cs166.1146/lectures/13/Small1 an explaination of Cuckoo Hashing, a randomized hash table scheme which, like 2-l hashing, has O(1) query time, but also has expected O(1) insertion time.
7. 9/26	Tue	Finish up Bloom filter analysis.	Slides. (./slides/lecture7/lecture7Annotated.pdf) Compressed slides. (./slides/lecture7/lecture7Compressed.pdf) Reading: Chapter 4 (http://infolab.stanford.edu/~ullman/mmds/ch4.pdf) of *Mining of Massive Datasets*, content on bloom filters. See here (http://cglab.ca/~morin/publications/ds/bloom-submitted.pdf) for full Bloom filter analysis.
8. 9/28	Thu	Min-Hashing for Distinct Elements. The median trick.	Slides. (./slides/lecture8/lecture8Annotated.pdf) Compressed slides. (./slides/lecture8/lecture8Compressed.pdf) Reading: Chapter 4 (http://infolab.stanford.edu/~ullman/mmds/ch4.pdf) of *Mining of Massive Datasets*, content on distinct elements counting.
9. 10/3	Tue	Distinct elements in pratice: Flajolet-Martin and HyperLogLog. Start on Jaccard similarity and motivation for the fast similarity search problem.	Slides. (./slides/lecture9/lecture9Annotated.pdf) Compressed slides. (./slides/lecture9/lecture9Compressed.pdf) Reading: The 2007 paper (http://algo.inria.fr/flajolet/Publications/FlFuGaMe07.pdf) introducing the popula HyperLogLog distinct elements algorithm. Chapter 3 (http://infolab.stanford.edu/~ullman/mmds/ch3.pdf) of *Mining of Massive Datasets*, content on Jaccard similarity, MinHash, and locality sensitive hashing.
10. 10/5	Thu	Fast similarity search via locality sensitive hashing. MinHashing for Jaccard similarity.	Slides. (./slides/lecture10/lecture10Annotated.pdf) Compressed slides. (./slides/lecture10/lecture10Compressed.pdf) Reading: Chapter 3 (http://infolab.stanford.edu/~ullman/mmds/ch3.pdf) of *Mining of Massive Datasets*, content on Jaccard similarity, MinHash, and locality sensitive hashing.
10/10	Tue	No Class. Monday class schedule followed.	
11. 10/12	Thu	Frequent elements estimation via Count-min sketch.	Slides (./slides/lecture11/lecture11Annotated.pdf). Compressed slides. (./slides/lecture11/lecture11Compressed.pdf) Reading: Notes (https://www.cs.dartmouth.edu/~ac/Teach/CS49-Fall11/Notes/lecnotes.pdf) (Ami Chakrabarti at Dartmouth) on streaming algorithms. See Chapters 2 and 4 for frequ elements. Some more notes (https://people.csail.mit.edu/rrw/6.045-2019/encalgs- the frequent elements problem. A website (https://sites.google.com/site/countmin with lots of resources, implementations, and example applications of count-min ske Algebra Review: Khan academy (https://www.khanacademy.org/math/linear-algeb
12. 10/17	Tue	Dimensionality reduction, low-distortion embeddings, and the Johnson Lindenstrauss Lemma.	Slides. (./slides/lecture12/lecture12Annotated.pdf) Compressed slides. (./slides/lecture12/lecture12Compressed.pdf) Reading: Chapter 2.7 (https://www.cs.cornell.edu/jeh/book.pdf) of *Foundations of Data Science* on the Joh Lindenstrauss lemma. Notes on the JL-Lemma (https://www.cs.cmu.edu/~avrim/Randalgs11/lectures/lect0314.pdf) (Anupam Gu Sparse random projections (https://arxiv.org/pdf/1004.4240.pdf) which can be mul more quickly.
13. 10/19	Thu	Midterm Review.	Slides. (./slides/lecture13/lecture13Annotated.pdf)
10/24	Tue	Midterm (In Class)	Study guide and review questions. (midtermConcepts.pdf)
			Spectral Methods
14. 10/26	Thu	Finish up the JL	Slides. (./slides/lecture14/lecture14Annotated.pdf) Compressed slides.

		Lemma proof. Example application to clustering.	(./slides/lecture14/lecture14Compressed.pdf) Reading: Chapter 2.7 (https://www.cs.cornell.edu/jeh/book.pdf) of *Foundations of Data Science* on the Joh Lindenstrauss lemma. Notes on the JL-Lemma (https://www.cs.cmu.edu/~avrim/Randalgs11/lectures/lect0314.pdf) (Anupam Gu
15. 10/31	Tue	Intro to principal component analysis, low-rank approximation, data-dependent dimensionality reduction. Orthogonal bases and projection matrices.	Slides (./slides/lecture15/lecture15Annotated.pdf). Compressed slides. (./slides/lecture15/lecture15Compressed.pdf) Reading: Chapter 3 (https://www.cs.cornell.edu/jeh/book.pdf) of *Foundations of Data Science* and Chapt (http://infolab.stanford.edu/~ullman/mmds/ch11.pdf) of *Mining of Massive Datasets* rank approximation and the SVD. Some good videos for linear algebra review. (https://www.3blue1brown.com/topics/linear-algebra) Some other good videos ov the SVD and related topics (https://www.youtube.com/watch? v=gXbThCXjZFM&list=PLMrJAkhIeNNSVjnsviglFoY2nXildDCcv&ab_channel=Ste (like orthogonal projection and low-rank approximation).
16. 11/02	Thu	Dual column/row view of low-rank approximation. Best fit subspaces and optimal low-rank approximation via eigendecomposition.	Slides (./slides/lecture16/lecture16Annotated.pdf). Compressed slides (./slides/lecture16/lecture16Compressed.pdf). Reading: Proof that optimal low-ran approximation can be found greedily (https://www.cs.cmu.edu/~venkatg/teaching/ infoage/book-chapter-4.pdf) (see Section 1.1). Chapter 3 (https://www.cs.cornell.edu/jeh/book.pdf) of *Foundations of Data Science* and Chapt (http://infolab.stanford.edu/~ullman/mmds/ch11.pdf) of *Mining of Massive Datasets* rank approximation.
17. 11/07	Tue	Finish up optimal low-rank approximation via eigendecomposition. Eigenvalues as a measure of low-rank approximation error.	Slides (./slides/lecture17/lecture17Annotated.pdf). Compressed slides (./slides/lecture17/lecture17Compressed.pdf). Reading: Chapter 3 (https://www.cs.cornell.edu/jeh/book.pdf) of *Foundations of Data Science* and Chapt (http://infolab.stanford.edu/~ullman/mmds/ch11.pdf) of *Mining of Massive Datasets* rank approximation.
18. 11/09	Thu	The singular value decomposition and connections to low-rank approximation. Applications of low-rank approximation beyond compression. Matrix completion and entity embeddings.	Slides (./slides/lecture18/lecture18Annotated.pdf). Compressed slides (./slides/lecture18/lecture18Compressed.pdf). Reading: Notes on SVD and its conn eigendecomposition/PCA (Roughgarden and Valiant at Stanford) (http://web.stanford.edu/class/cs168/l/l9.pdf). Notes on matrix completion (http://people.seas.harvard.edu/~minilek/cs229r/fall15/lec/lec22.pdf), with proof c under incoherence assumptions (Jelani Nelson at Harvard). Levy Goldberg paper (https://papers.nips.cc/paper/5477-neural-word-embedding-as-implicit-matrix-factorization.pdf) on word embeddings as implicit low-rank approximation.
19. 11/14	Tue	Spectral graph theory and spectral clustering.	Slides (./slides/lecture19/lecture19Annotated.pdf). Compressed slides (./slides/lecture19/lecture19Compressed.pdf). Reading: Chapter 10.4 (http://infolab.stanford.edu/~ullman/mmds/ch10.pdf) of *Mining of Massive Datasets* spectral graph partitioning. For a lot more interesting material on spectral graph me Dan Spielman's lecture notes (http://www.cs.yale.edu/homes/spielman/561/). Grea spectral graph methods (http://web.stanford.edu/class/cs168/l/l11.pdf) (Roughgar Valiant at Stanford).
20. 11/16	Thu	The stochastic block model.	Slides (./slides/lecture20/lecture20Annotated.pdf). Compressed slides (./slides/lecture20/lecture20Compressed.pdf). Reading: Dan Spielman's lecture not stochastic block model, including matrix concentration + David-Kahan perturbatio (https://www.cs.yale.edu/homes/spielman/561/lect21-15.pdf). Further stochastic model notes (http://www.stat.cmu.edu/~arinaldo/Teaching/36710/F18/Scribed_Lectures/Nov1 (Alessandro Rinaldo at CMU). A survey (http://www.princeton.edu/~eabbe/publications/abbe_FNT_2.pdf) of the vast litera the stochastic block model, beyond the spectral methods discussed in class (Emmar at Princeton).

21. 11/21	Tue	Computing the SVD: power method. Krylov methods. Bonus material: connection to random walks and Markov chains.	Slides (./slides/lecture21/lecture21Annotated.pdf). Compressed slides (./slides/lecture21/lecture21Compressed.pdf). Reading: Chapter 3.7 (https://www.cs.cornell.edu/jeh/book.pdf) of *Foundations of Data Science* on the pov for SVD. Some notes on the power method. (http://web.stanford.edu/class/cs168/l (Roughgarden and Valiant at Stanford).
11/23	Thu	No Class. Thanksgiving recess.	
11/28	Tue	No Class. Professor traveling.	
		Optimization	
22. 11/30	Thu	Start on optimization and gradient descent.	Slides (./slides/lecture22/lecture22Annotated.pdf). Compressed slides (./slides/lecture22/lecture22Compressed.pdf). Reading: Chapters I and III of these (https://ee227c.github.io/notes/ee227c-notes.pdf) (Hardt at Berkeley). Multivarial review, e.g., through: Khan academy (https://www.khanacademy.org/math/multiva calculus/multivariable-derivatives)
23. 12/05	Tue	Gradient descent analysis for convex Lipschitz functions.	Slides (./slides/lecture23/lecture23Annotated.pdf). Compressed slides (./slides/lecture23/lecture23Compressed.pdf). Reading: Chapters I and III of these (https://ee227c.github.io/notes/ee227c-notes.pdf) (Hardt at Berkeley).
24. 12/07	Thu	Constrained optimization and projected gradient descent. Course conclusion/review.	Slides (./slides/lecture24/lecture24Annotated.pdf). Compressed slides (./slides/lecture24/lecture24Compressed.pdf).
12/14, 10:30am - 12:30pm		Final Exam.	Study guide and review questions. (finalConcepts.pdf)