خلاصه‌سازی برای مه‌داده

ترم پاییز ۱۳۹۹–۱۴۰۰

جلسه اول

خلاصه‌سازی برای مه‌داده

توصیف

درس

خلاصه‌سازی مه‌داده

- مه‌داده، کلان‌داده

- پیش‌نیاز:
  - الگوریتم، احتمال

مباحث درس

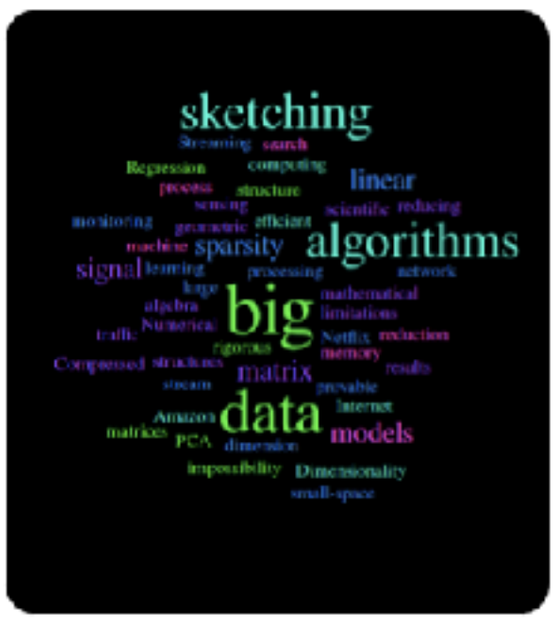
- ★ Topic 1: Sketching
- ★ Topic 2: Streaming
- ★ Topic 3: Dimensionality Reduction
- ★ Topic 4: Large-Scale Matrix Computations
- ★ Topic 5: Compressed Sensing
- ★ Topic 6: Sparse Fourier Transform

منابع

sketchingbigdata.org

SKETCHING ALGORITHMS

Fall 2020



Current offering (Fall 2020)

- Sketching Algorithms

Related previous offerings

- Sketching Algorithms for Big Data: Piotr Indyk (MIT), Jelani Nelson (Harvard).
- Sketching, Streaming, and Sub-linear Space algorithms: Piotr Indyk (MIT).
- Algorithms for Big Data: Jelani Nelson (Harvard).

Related courses at other schools

- Algorithmic Techniques for Big Data: Moses Charikar (Stanford).
- Algorithmic Techniques for Big Data Analysis: Barna Saha (University of Minnesota Twin Cities).

نسخه ۲۰۲۰

Jelani Nelson

برنامه نسخه ۲۰۲۰

Lec	Date	Topic	Readin
1	Wednesday,	logistics, course topics, approximate	§ 2.1
2	Monday,	distinct elements, k-wise	§
3	Wednesday,	geometric sampling	§ 2.2.3
	Monday,	**NO CLASS**	
4	Wednesday,	quantiles	§ 2.3
5	Monday,	lower bounds via compression	§ 3.1
6	Wednesday,	lower bounds via communication	§ 3.2

نسخه ۲۰۱۷

نسخه ۲۰۱۷

Lec#	Notes
1	logistics, course topics, basic tail bounds (Markov,
2	distinct elements
3	turnstile streaming, $\ell 2$ norm estimation via AMS sketch,
4	ℓp norm estimation
5	Johnson-Lindenstrauss lower bound
6	JL lower bound wrap-up, heavy hitters
7	CountMin sketch, CountSketch, insertion-only
8	continuous monitoring, chaining
9	Fast Johnson-Lindenstrauss Transform
10	Sparse Johnson-Lindenstrauss Transform
11	randomized linear algebra: approximate matrix multiplication,

12	randomized linear algebra: subspace embeddings, regression,
13	Compressive sensing, RIP, L1 minimization
14	Iterative hard thresholding, model-based compressive
15	Sparse recovery using sparse matrices, RIP1
16	Sparse Fourier Transform
17	Sparse Fourier Transform (continued)
18	ℓp sampling, graph sketching
19	Streaming algorithms for geometric problems
20	lower bounds
21	Streaming algorithms for geometric problems
22	**Guest Lecture**: Cameron Musco
23	**Guest Lecture**: Noga Alon
24	**Guest Lecture**: Christopher Musco

کلاس را چگونه می‌گذرانیم؟

- کوئیزک ساده کلاسی
  - سازوکار مناسب
  - سوال‌های سلیقه‌ای و سوال‌های خیلی ساده
- سوال –> دست بلند کردن –>
- کوئیز (با اعلام قبلی) کتاب_باز
- تمرین: نظری

- نمره‌دهی
  - کوئیزک: ۳ نمره (از ۶ نمره)
  - کوئیز: ۲
  - میان‌ترم: ۴
  - پایان‌ترم: ۵
  - تمرین: ۳
    - بدون تاخیر
    - ۳ تا تمرین تصحیح نمی‌شود. مشابهش در کوئیز و میان و پایان‌ترم. عملی تصحیحی.
  - جزوه: ۱
  - ارائه + گزارش: ۲ نمره
- راستی‌آزمایی نهایی در صورت تشخیص تقلب

سمینار همراه درس

- برای همه نیست!