# Overview

Daniel Hsu

COMS 4772

# About COMS 4772

- "Advanced machine learning"

# About COMS 4772

- "Advanced machine learning"
  - But actually: "machine learning theory"

# About COMS 4772

- "Advanced machine learning"
  - But actually: "machine learning theory"
- Website:
  `http://www.cs.columbia.edu/~djhsu/coms4772-f16/`

# About COMS 4772

- "Advanced machine learning"
  - But actually: "machine learning theory"
- Website: `http://www.cs.columbia.edu/~djhsu/coms4772-f16/`
  - Course information, policies, academic rules of conduct, etc.

# About COMS 4772

- "Advanced machine learning"
  - But actually: "machine learning theory"
- Website:
  `http://www.cs.columbia.edu/~djhsu/coms4772-f16/`
  - Course information, policies, academic rules of conduct, etc.
- Courseworks, Piazza: links on website

# About you

- Satisfy all prerequisites:

# About you

- Satisfy all prerequisites:
  - “machine learning”

# About you

- Satisfy all prerequisites:
  - “machine learning”
  - multivariate calculus

# About you

- Satisfy all prerequisites:
  - “machine learning”
  - multivariate calculus
  - linear algebra

# About you

- Satisfy all prerequisites:
    - "machine learning"
    - multivariate calculus
    - linear algebra
    - probability theory

# About you

- Satisfy all prerequisites:
  - "machine learning"
  - multivariate calculus
  - linear algebra
  - probability theory
  - algorithms

# About you

- Satisfy all prerequisites:
  - “machine learning”
  - multivariate calculus
  - linear algebra
  - probability theory
  - algorithms
  - mathematical maturity

# About you

- Satisfy all prerequisites:
  - “machine learning”
  - multivariate calculus
  - linear algebra
  - probability theory
  - algorithms
  - mathematical maturity
  - can read/write mathematical arguments, derivations, and proofs

# About you

- Satisfy all prerequisites:
  - “machine learning”
  - multivariate calculus
  - linear algebra
  - probability theory
  - algorithms
  - mathematical maturity
  - can read/write mathematical arguments, derivations, and proofs
- You have until next lecture to “page in” these topics.

# About you

- Satisfy all prerequisites:
  - "machine learning"
  - multivariate calculus
  - linear algebra
  - probability theory
  - algorithms
  - mathematical maturity
  - can read/write mathematical arguments, derivations, and proofs
- You have until next lecture to "page in" these topics.
  - See Homework 0

# About you (relative to this class)

- Read papers/notes (posted on website)

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets ($\sim$4)

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets ($\sim$4)
  - Write them up in $\LaTeX$or something of similar quality

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets ($\sim$4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets (∼4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project
  - E.g., new, interesting theoretical result

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets (~4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project
  - E.g., new, interesting theoretical result
  - E.g., simplify an existing, complex result in a non-trivial way

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets (∼4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project
  - E.g., new, interesting theoretical result
  - E.g., simplify an existing, complex result in a non-trivial way
  - E.g., *high quality* survey paper that unifies several papers

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets (∼4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project
  - E.g., new, interesting theoretical result
  - E.g., simplify an existing, complex result in a non-trivial way
  - E.g., *high quality* survey paper that unifies several papers
  - Cannot "just" implement an algorithm and run some experiments

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets (~4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project
  - E.g., new, interesting theoretical result
  - E.g., simplify an existing, complex result in a non-trivial way
  - E.g., *high quality* survey paper that unifies several papers
  - Cannot "just" implement an algorithm and run some experiments
  - Project report / presentation (possibly a poster session) at end of semester (maybe during "final exam" time)

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets (~4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project
  - E.g., new, interesting theoretical result
  - E.g., simplify an existing, complex result in a non-trivial way
  - E.g., *high quality* survey paper that unifies several papers
  - Cannot "just" implement an algorithm and run some experiments
  - Project report / presentation (possibly a poster session) at end of semester (maybe during "final exam" time)
- Abide by course policies, academic rules of conduct

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets ($\sim$4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project
  - E.g., new, interesting theoretical result
  - E.g., simplify an existing, complex result in a non-trivial way
  - E.g., *high quality* survey paper that unifies several papers
  - Cannot "just" implement an algorithm and run some experiments
  - Project report / presentation (possibly a poster session) at end of semester (maybe during "final exam" time)
- Abide by course policies, academic rules of conduct
  - See website

# About you (relative to this class)

- Read papers/notes (posted on website)
- Attend lectures (mostly at blackboard, sometimes slides)
- Solve problem sets (~4)
  - Write them up in LaTeXor something of similar quality
- Work on theoretical research project
  - E.g., new, interesting theoretical result
  - E.g., simplify an existing, complex result in a non-trivial way
  - E.g., *high quality* survey paper that unifies several papers
  - Cannot "just" implement an algorithm and run some experiments
  - Project report / presentation (possibly a poster session) at end of semester (maybe during "final exam" time)
- Abide by course policies, academic rules of conduct
  - See website
  - Violators reported to the Dean's office, get failing grade for assignment and/or course

# About the course staff

- Instructor: Prof. Daniel Hsu

# About the course staff

- Instructor: Prof. Daniel Hsu
  - Website: `http://www.cs.columbia.edu/~djhsu/`

# About the course staff

- Instructor: Prof. Daniel Hsu
  - Website: `http://www.cs.columbia.edu/~djhsu/`
  - Research in algorithmic statistics, machine learning

# About the course staff

- Instructor: Prof. Daniel Hsu
  - Website: `http://www.cs.columbia.edu/~djhsu/`
  - Research in algorithmic statistics, machine learning
- Course assistants: Rob and Mark

## About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., "subpopulations" in genetic study panel

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., “subpopulations” in genetic study panel
  - E.g., “communities” in a social network

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., "subpopulations" in genetic study panel
  - E.g., "communities" in a social network
  - E.g., "topics" in a corpus of documents

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., "subpopulations" in genetic study panel
  - E.g., "communities" in a social network
  - E.g., "topics" in a corpus of documents
  - Many omissions (e.g., PAC learning, Bayesian analysis)

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., "subpopulations" in genetic study panel
  - E.g., "communities" in a social network
  - E.g., "topics" in a corpus of documents
  - Many omissions (e.g., PAC learning, Bayesian analysis)
- Role of theoretical analysis in machine learning

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., "subpopulations" in genetic study panel
  - E.g., "communities" in a social network
  - E.g., "topics" in a corpus of documents
  - Many omissions (e.g., PAC learning, Bayesian analysis)
- Role of theoretical analysis in machine learning
  - Beyond worst-case analysis: also have model of "input" (data)

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., "subpopulations" in genetic study panel
  - E.g., "communities" in a social network
  - E.g., "topics" in a corpus of documents
  - Many omissions (e.g., PAC learning, Bayesian analysis)
- Role of theoretical analysis in machine learning
  - Beyond worst-case analysis: also have model of "input" (data)
  - *Best case analysis*, but assumptions usually violated in practice

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., "subpopulations" in genetic study panel
  - E.g., "communities" in a social network
  - E.g., "topics" in a corpus of documents
  - Many omissions (e.g., PAC learning, Bayesian analysis)
- Role of theoretical analysis in machine learning
  - Beyond worst-case analysis: also have model of "input" (data)
  - *Best case analysis*, but assumptions usually violated in practice
  - Often lags practice, but not always (e.g., boosting, $k$-means++)

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., “subpopulations” in genetic study panel
  - E.g., “communities” in a social network
  - E.g., “topics” in a corpus of documents
  - Many omissions (e.g., PAC learning, Bayesian analysis)
- Role of theoretical analysis in machine learning
  - Beyond worst-case analysis: also have model of “input” (data)
  - *Best case analysis*, but assumptions usually violated in practice
  - Often lags practice, but not always (e.g., boosting, $k$-means++)
  - Framework for reasoning about machine learning algorithms

# About COMS 4772 (again)

- Techniques for designing/analyzing machine learning algorithms
  - Focus on simple statistical models of data
  - E.g., "subpopulations" in genetic study panel
  - E.g., "communities" in a social network
  - E.g., "topics" in a corpus of documents
  - Many omissions (e.g., PAC learning, Bayesian analysis)
- Role of theoretical analysis in machine learning
  - Beyond worst-case analysis: also have model of "input" (data)
  - *Best case analysis*, but assumptions usually violated in practice
  - Often lags practice, but not always (e.g., boosting, $k$-means++)
  - Framework for reasoning about machine learning algorithms
  - Suggest new algorithmic techniques

# About COMS 4772 (tentative list of topics)

1. High-dimensional data

# About COMS 4772 (tentative list of topics)

1. High-dimensional data
   - concentration of measure, random linear maps

# About COMS 4772 (tentative list of topics)

1. High-dimensional data
    - concentration of measure, random linear maps
    - applications: least squares regression, $k$-means clustering, Gaussian mixtures

# About COMS 4772 (tentative list of topics)

1. High-dimensional data
   - concentration of measure, random linear maps
   - applications: least squares regression, $k$-means clustering, Gaussian mixtures
2. Low-rank matrix approximation

# About COMS 4772 (tentative list of topics)

1. High-dimensional data
   - concentration of measure, random linear maps
   - applications: least squares regression, $k$-means clustering, Gaussian mixtures
2. Low-rank matrix approximation
   - PCA, SVD, NMF, power iteration

# About COMS 4772 (tentative list of topics)

1. High-dimensional data
   - concentration of measure, random linear maps
   - applications: least squares regression, $k$-means clustering, Gaussian mixtures
2. Low-rank matrix approximation
   - PCA, SVD, NMF, power iteration
   - applications: Gaussian mixtures, $k$-means clustering, planted partition models, topic models

# About COMS 4772 (tentative list of topics)

1. High-dimensional data
   - concentration of measure, random linear maps
   - applications: least squares regression, $k$-means clustering, Gaussian mixtures
2. Low-rank matrix approximation
   - PCA, SVD, NMF, power iteration
   - applications: Gaussian mixtures, $k$-means clustering, planted partition models, topic models
3. Higher-order interactions

# About COMS 4772 (tentative list of topics)

1. High-dimensional data
    - concentration of measure, random linear maps
    - applications: least squares regression, $k$-means clustering, Gaussian mixtures
2. Low-rank matrix approximation
    - PCA, SVD, NMF, power iteration
    - applications: Gaussian mixtures, $k$-means clustering, planted partition models, topic models
3. Higher-order interactions
    - higher-order tensors, tensor decompositions, power iteration

# About COMS 4772 (tentative list of topics)

1. High-dimensional data
   - concentration of measure, random linear maps
   - applications: least squares regression, $k$-means clustering, Gaussian mixtures
2. Low-rank matrix approximation
   - PCA, SVD, NMF, power iteration
   - applications: Gaussian mixtures, $k$-means clustering, planted partition models, topic models
3. Higher-order interactions
   - higher-order tensors, tensor decompositions, power iteration
   - applications: Gaussian mixtures, ICA, latent Dirichlet allocation