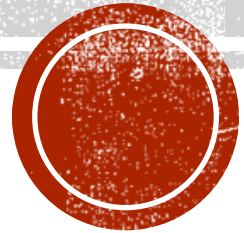
MACHINE LEARNING
CS 229 / STATS 229



INTRODUCTION

Anand Avati

Teaching Assistants

Carolyn Kim

Chenru (Helen) Liu

Hojat Ghorbani

Jian Zhang

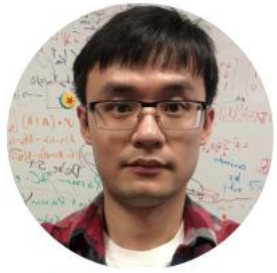
Nandita Bhaskhar

Ruiyang Song

Sandeep Chinchali

Shaimaa Bakr

Soyeon Jung



Zhenglin Geng

GOALS

By the end of the course,

- Be an expert in ML (understand the internals of ML algorithms)
- Be able to build ML applications (know which algorithms to use when)
- Be able to start ML research (read research papers)

PREREQUISITES

- Basic computer science principles
  - Big-O notation
  - Comfortably write non-trivial code in Python/numpy
- Probability (CS 109, STATS 116 etc.)
  - Random Variables
  - Expectations
  - Distributions
- Linear Algebra & Multivariate/Matrix Calculus (MATH 51, etc.)
  - Gradients and Hessians
  - Eigenvalue/vector

HONOR CODE

- Form study groups (strongly encouraged!)
- Independently write-up homework and code
- **It is an honor code violation to intentionally refer to a previous year's assignments.**

COURSE STRUCTURE

- 3 Homeworks : 60% (3 x 20%)
- Final exam (take home) : 40%

LOGISTICS

- Course website (calendar, deadlines, notes) - http://cs229.stanford.edu
- Piazza
- Gradescope

WHAT IS ML?

- Term "Machine Learning" coined by Arthur Samuel in 1959.
  - Samuel Checkers-playing Program

- Common definition (by Tom Mitchell):
  - ***Machine Learning is the study of computer algorithms that improve automatically through experience***

- Subfield of Artificial Intelligence (AI)
  - The hottest subfield - reinvigorated interest in AI due to deep learning!

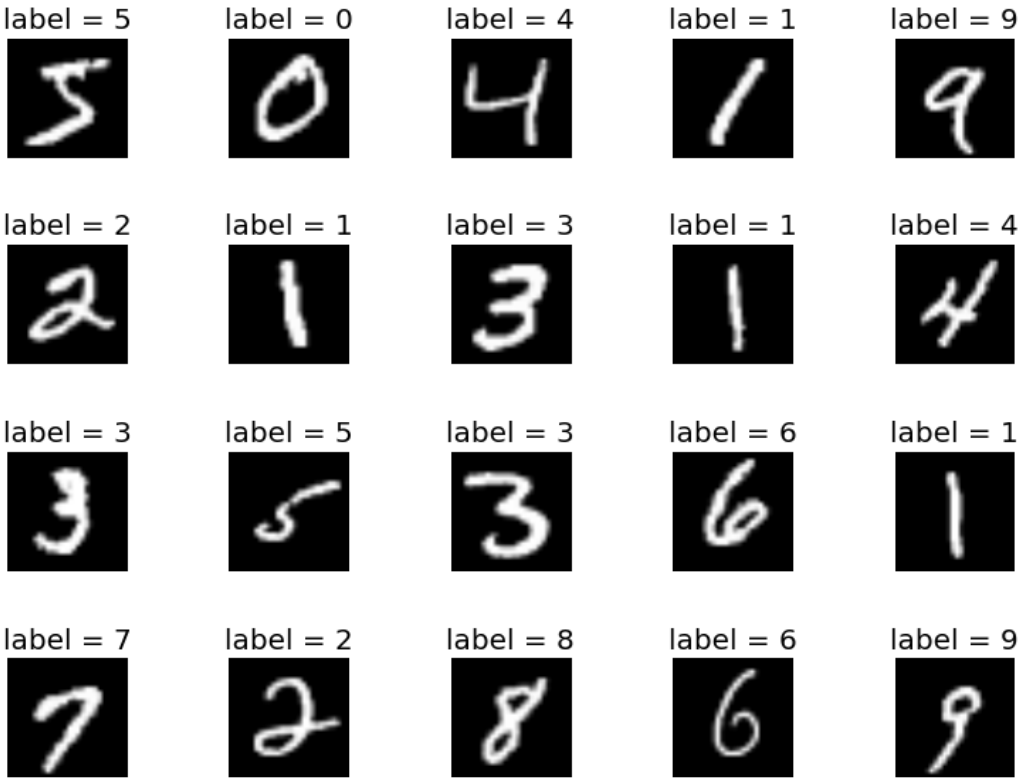
RECENT PROGRESS

- Computer Vision / Image Recognition
  - ImageNet
  - Convolutional Neural Networks
  - Autonomous driving
- Speech Recognition
  - Voice assistants
- Language Translation
  - Google Translate
  - Unsupervised Translation
- Game Playing / Deep Reinforcement Learning
  - ATARI
  - AlphaGo

COURSE PREVIEW

- Supervised learning [Input -> Output mapping]
  - Classification vs Regression
  - Parametric vs Non-parametric
  - Generative vs Discriminative
  - Probabilistic vs Non-probabilistic
- Unsupervised learning [Learn interesting structures in the data]
  - Clusters vs Subspaces
  - Probabilistic vs Non-probabilistic
- Deep Learning [Learning representations]
  - Our focus: supervised setting
- Learning Theory
  - Bias-Variance Tradeoff
  - Generalization and Uniform Convergence
- Reinforcement Learning [Sequential Decision Making]

PREVIEW - SUPERVISED



PREVIEW – UNSUPERVISED (ICA)

Mixed

Separated

PREVIEW – UNSUPERVISED (ICA - II)

Mixed

Separated

PREVIEW – REINFORCEMENT LEARNING