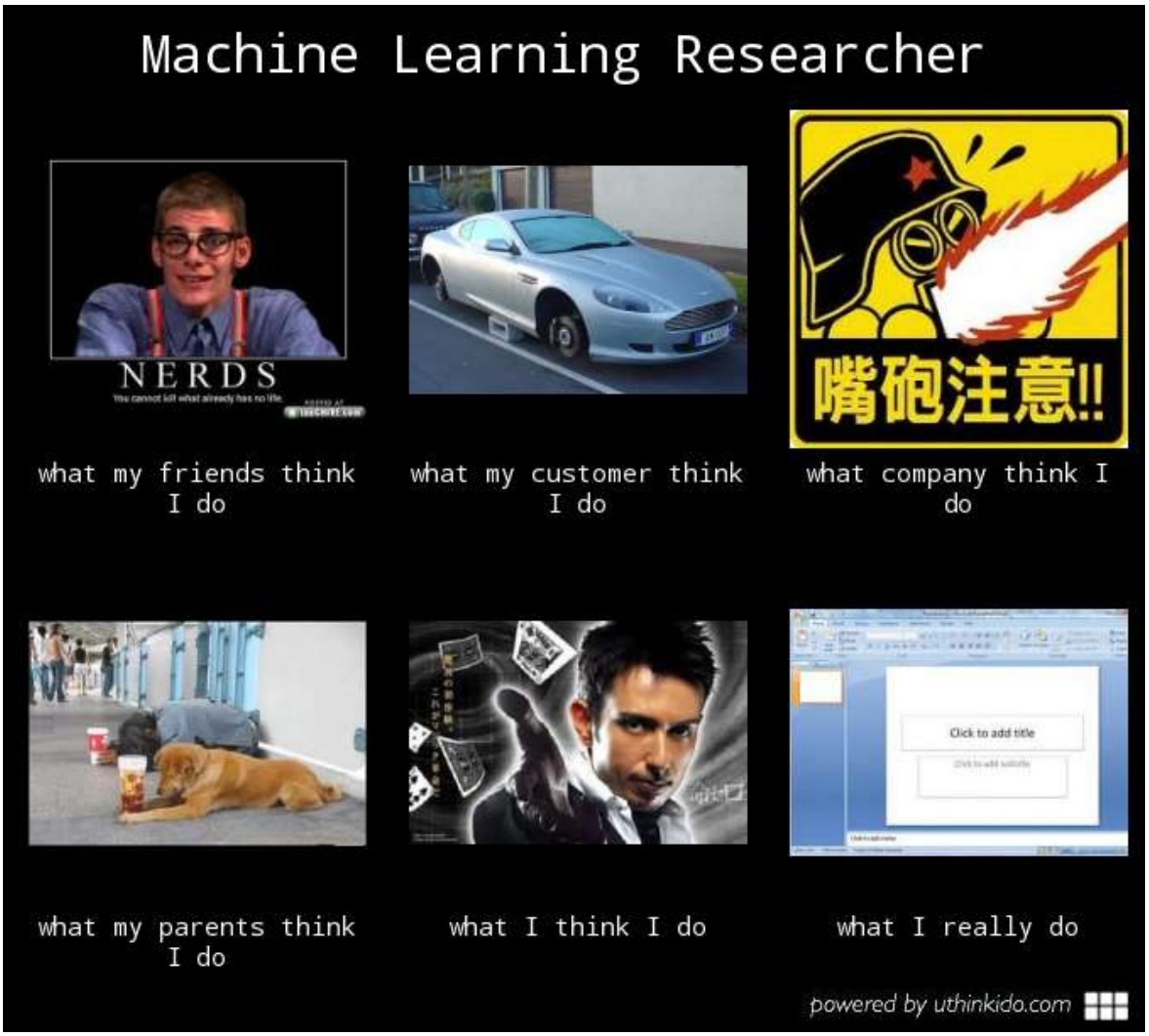
CA684 Jia Yuan Yu

1: Intro

DCU February 13, 2014

1 What is ML?



- A branch of AI.
- A combination of statistics (guarantees on error), control (prediction, estimation, filtering), optimization, etc.
- A precise science.

1.1 Wikipedia definition

`Tom M. Mitchell provided a widely quoted, more formal definition: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E".`

Example: Learning the mean of a random variable from observations.

Theorem 1.1 (Hoeffding)**.** *Let $X_1, X_2, \ldots$ be i.i.d. random variables that take values in the interval $[a, b]$, and have mean μ. Let*

$$\hat{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Then, for every n and $\varepsilon > 0$:

$$\mathbb{P}\left(\left|\hat{X}_n - \mu\right| \geq \varepsilon\right) \leq 2 \exp\left(-\frac{2n\varepsilon^2}{b - a^2}\right).$$

1.2 What is it good for?

- Pattern recognition, classification.
- Decision-making.
- Getting a job.

2 What are we going to learn?

- Supervised learning: ability to generalize from training data to unseen data.
  - Linear discrimination, nearest-neighbor rules, SVM
  - Consistency, VC theory
  - Maximum likelihood,
  - Regression, Regularization
  - Neural networks, feature extraction
  - Computer vision
  - Deep learning
- Online learning: sequential decision-making
  - Stochastic setting
  - Adversarial setting
  - Full-information (Combining expert-advice)

  - Partial-information (Bandit problem)
- Markov decision problems
  - Reinforcement learning
  - Stochastic approximation

3 What won't be covered?

- Unsupervised learning: clustering.
  - However, you can take CA683!
- Ensemble methods (boosting).
- Dimension reduction: PCA, dictionary learning.
- Natural language processing (you can read about it).
- Graphical models, Bayesian networks.
- Learning in games, against opponents (think poker).
- Robotics.
- Cognitive science.
- Logic.
- Anomaly detection.

4 What are the rules?

- Final exam: 75%
- Homeworks, projects: 25%

Project options:

- Read about a topic in ML and present the class,
- Programming: apply techniques from class on real data sets.

Project evaluation: Peer-review on last two lectures.

5 Reading material

Books:

- A Probabilistic Theory of Pattern Recognition
- Prediction, Learning, and Games
- Markov Decision Processes: Approximate dynamic programming, adaptive control, models, theory and computation

If hungry for more:

- The Elements of Statistical Learning.
- Machine Learning: a Probabilistic Perspective