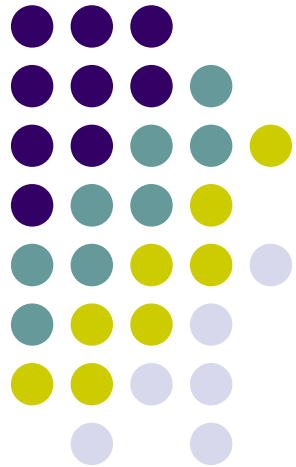
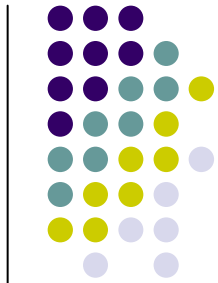
Machine Learning: Introduction

Zhang Hongxin
zhx@cad.zju.edu.cn

State Key Lab of CAD&CG, ZJU
2008-02-21



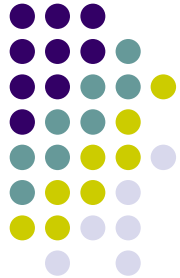
Syllabus



- Covers a set of Machine learning techniques –from basic and state-of-the-art.
- You will learn:
  - PCA, MDS,K-mean, spectrum based clustering, Naïve-Bayes classification, boosting, logistic regression, decision tree, EM, HMM, Kalman filtering…
- Tell the stories behind the algorithms, theory and applications.

- It is going to be fun and hard work.

Rough schedule



- 02.21: Introduction
- 02.28: Classification
- 03.06: Clustering
- 03.13: HMM & Kalman Filtering

Principle

- Simple is beauty!

- Make a balance between theories and real applications

Final report

- 每个短学期提交一个课程报告
- 截止日期：
  - 春季学期：**2008/06/01**
  - 夏季学期：**2008/09/01**
  - 及早提交，及早获取成绩！
- 评价：
  - 基本评价：**80%**（翻译，理解，自己的体会）
  - 附加评价：**20%**（提交论文的算法实现，附程序，最好能有源程序或者核心代码片段）

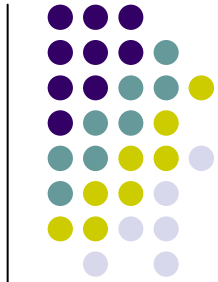
Final report

- Paper reading report.
  - Read a typical learning paper. Report the main idea and your own opinions.
    - Paper source:
      - ACM SIGGRAPH / EUGROGRAPHICS / SCA / Pacific Graphics,
      - ICCV / ECCV / ACCV, CVPR,
      - NIPS / ICML / IJCAI / UAI / AAAI
      - SIGMOD, SIGKDD

      - Corresponding Journals …
    - Report can be in English or Chinese. And please hand out in PDF format. （标注姓名，学号）

    - zhx@cad.zju.edu.cn **or** huawei@cad.zju.edu.cn
    - chenwei@cad.zju.edu.cn **or** hwlin@cad.zju.edu.cn

Prerequisites



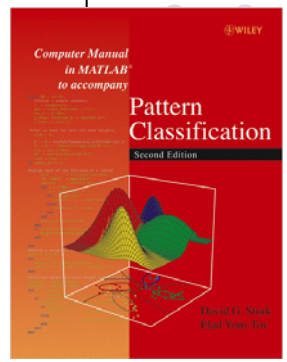
- Probabilities
  - Distributions, densities, marginalization…
- Basic statistics
  - Moments, typical distributions, regression…
- Algorithms
  - Dynamic programming, basic data structures, complexity…
- Programming
  - Mostly your choice of language: C/C++, MATLAB, JAVA
- We provide some background, but the class will be fast paced

- Ability to deal with “abstract mathematical concepts”

源自88上某位网友的签名档

- 鉴于大多数博士们在之后的生活中并没有从事博士生期间的课题的研究，甚至根本不再做研究工作，我想攻读博士的目标应该是:
  1. 成为一个身体强壮的人
  2. 成为一个意志强悍的人
  3. 成为一个能系统思考，从混沌的一堆问题中提炼主要的具体的问题的人
  4. 成为一个能解决具体问题的人

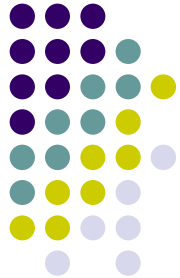
Text books



- Machine Learning
  - by Tom Mitchell
- Pattern Classification (2nd Edition)
  - by Duda, Hart and Stork
- Information Theory, Inference, and Learning Algorithms
  - by David MacKay
- **Statistical Inference**,
  - by George Casella and Roger L. Berger.
- And more …

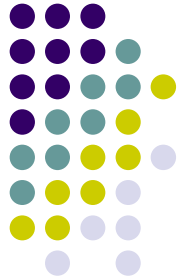
- **All above books are optional. Everyone have their own learning algorithms, ☺**

Internet resources



- http://www.cad.zju.edu.cn/home/zhx/ML/

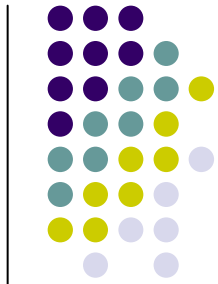
Enjoy!



- Machine Learning is becoming ubiquitous in science, engineering and beyond.
- This class should give you the basic foundation for applying ML and developing new methods.

- The fun begins…

Reference



- http://www-2.cs.cmu.edu/~guestrin/Class/10701/slides/CarlosIntro.pdf