



# 机器学习

## I. 介绍

# 自我介绍

➤ 孟德宇

➤ 数学与统计学院统计系

➤ Email:

[dymeng@mail.xjtu.edu.cn](mailto:dymeng@mail.xjtu.edu.cn)

➤ 个人主页:

<http://dymeng.gr.xjtu.edu.cn>

# 重要事项


- 上课时间：  
1-12周 周三1—2节
- 上课地点：中2-2250
- 组成学习小组：4-5人
- 我班共47人，期待形成10-12个小组

# 重要事项

## ➤ 考核方式：

➤ 程序实验报告：三份

➤ 科研学术报告：一份



与机器学习最相关  
的一个词是？

# 1. 介绍

- 人工智能简述
- 什么是机器学习?
- 机器学习历史
- 机器学习典型应用
- 典型机器学习问题
- 本课程计划讲授内容

# ● 人工智能诞生

- 1956年，“人工智能”这个词首次出现在一个持续2个月、只有10个人参加的研讨会提案上
- 达特茅斯夏季人工智慧研究计划（Dartmouth Summer Research Project on Artificial Intelligence）
- 参与人：
  - 约翰·麦卡锡
  - 马文·闵斯基
  - 克劳德·香农
  - 司马贺
  - 艾伦·纽厄尔
  - N. Rochester
  - Ray Solomonoff
  - Oliver Selfridge
  - Trenchard More
  - Arthur Samuel

# ● 人工智能三大学派

## ○ 符号主义学派

- 核心思想：认知即计算，通过对符号的演绎和逆演绎进行结果预测
- 巅峰作品：专家系统

## ○ 行为主义学派

- 核心思想：控制论及感知-动作型控制系统
- 巅峰作品：维纳滤波和演化算法

## ○ 连接主义（贝叶斯）学派

- 核心思想：神经网络及神经网络间的连接机制与学习算法
- 巅峰成果：支撑向量机，神经网络



# 人脸识别新技术准确率超99%：比肉眼更加精准

2014年06月23日 00:28 [新浪科技](#)  [微博](#) [收藏本文](#)

A<sup>-</sup> A<sup>+</sup>

## 头条推荐

- [【新闻】2016清明端午中秋可各拼出9天假](#)
- [【体育】世锦赛-丁俊晖夺冠 袖活！莱斯特英超夺冠](#)



人脸识别是计算机视觉和人工智能研究领域一个重要挑战，在公共安全、执法、移动互

联网和娱乐领域都有大量应用。它也成为检验人工智能是否可以解决其他特定智能问题上

8 [网信办牵头入驻调查医疗推广 百度周一早盘股价](#)

9 [停止开发低端Atom处理器 Intel终于放弃移动市](#)

10 [比特币之价现息 2.2个澳洲企业宣布向自己就是](#)

# 微软研究院喜报：电脑识图能力已超越人类

腾讯科技[微博] 晨曦 2015-02-16 11:33 我要分享

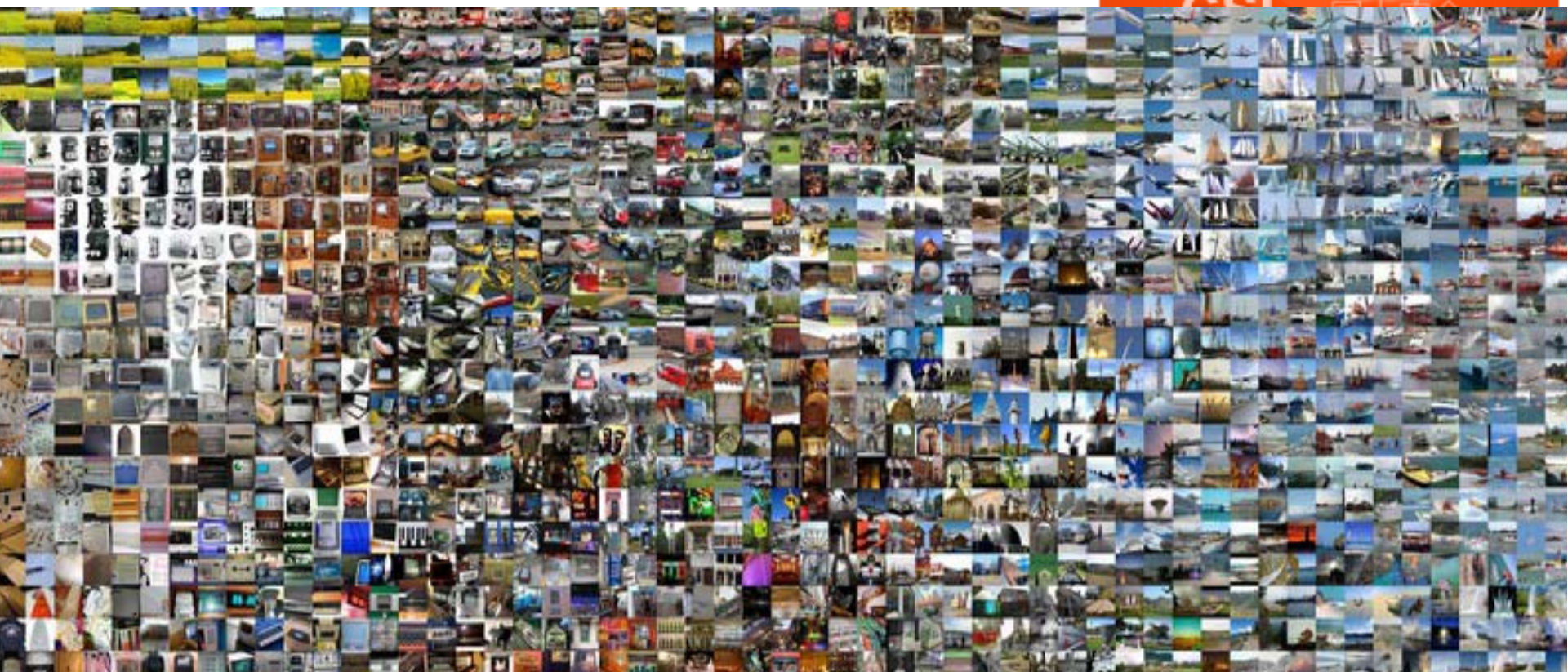
1



腾讯视频 美剧频道  
V.QQ.COM

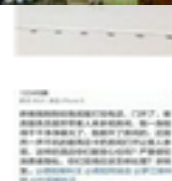
妙警贼探

欢乐合唱团



在移动互联网和物联网时代，图像识别技术越来越重要，比如手机中的图片管理软件，能够自动对照片中的人脸进行识别，将不同时间拍摄的照片进行归类，另外各种机器人也需要对周边事物进行快速的图像识别，帮助自己顺利行进。

日前，微软旗下的“微软研究院”公布了一个图片识别和计算机视觉技术的



遇险情 险机毁人亡

女房客正准备洗澡 德阳酒店服务员打开房门



攀枝花现3.4长巨型眼镜蛇 消防惊险活捉





李彦宏：今年成立百度首家研究院 专攻深度学习

2013-01-21 16:30:15 次阅读 次推荐 稿源

上周六，李彦宏在2012百度年会上发表演讲，再一次技术创新对于互联网的重要性，并宣布百度将在2013年历史上第一个研究院——Institute of Deep Learning (简称IDL)，初期将专注于Deep learning（深度学习）研究。



C 极客头条

10

Facebook成立人工智能实验室，大牛LeCun领军

Facebook

人工智能

Yann LeCun

深度学习



網易科技

网易首页 > 网易科技 > 互联网 > 正文

Twitter收购深度学习创业公司Madbits

2014-07-30 15:36:24 来源: 网易科技报道 有42人参与 分享到

CSDN首页 > 业界

谷歌收购DNNresearch，下一个帝国呼之欲出

发表于 2013-03-13 08:54 | 8745次阅读 | 来源 CSDN | 51 条评论 | 作者 张勇

Google

Geoffrey Hinton

机器学习

DNNresearch

**摘要：**近日Google收购了多伦多大学的研究公司，公司的负责人是Geoffrey Hinton，他是基于语音和图片上的搜索，这也使得谷歌的大帝国。

CSDN首页 > 云计算

【快讯】360建人工智能研究院 深度学习专家颜水成任院长及首席科学家

发表于 2015-09-28 10:47 | 1056次阅读 | 来源 CSDN | 1 条评论 | 作者 CSDN

深度学习

奇虎360

计算机视觉

AI

颜水成

大数据

# 1. 介绍

- 人工智能简述
- 什么是机器学习?
- 机器学习历史
- 机器学习典型应用
- 典型机器学习问题
- 本课程计划讲授内容



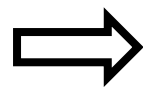
# 什么是机器学习

要学什么

从哪里学

怎样学习

要学什么



决策函数



物体识别



{狗, 猫, 马, 花, ...}

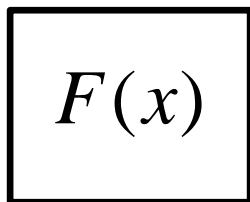


天气检测



雾霾的概率多大

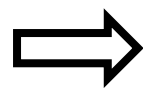
$x$



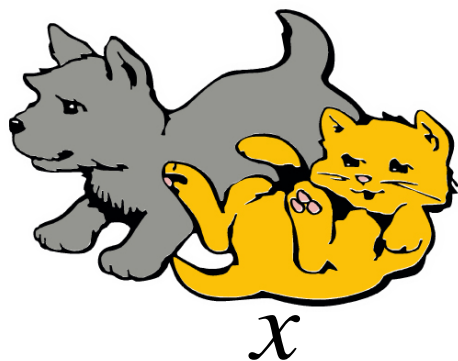
$y$

$\left\{ \begin{array}{l} \text{类标签} \\ \text{(分类问题)} \\ \\ \text{实数} \\ \text{(估计/回归问题)} \end{array} \right.$

从哪里学



训练数据



$F(x)$

{狗, 猫}

$y$

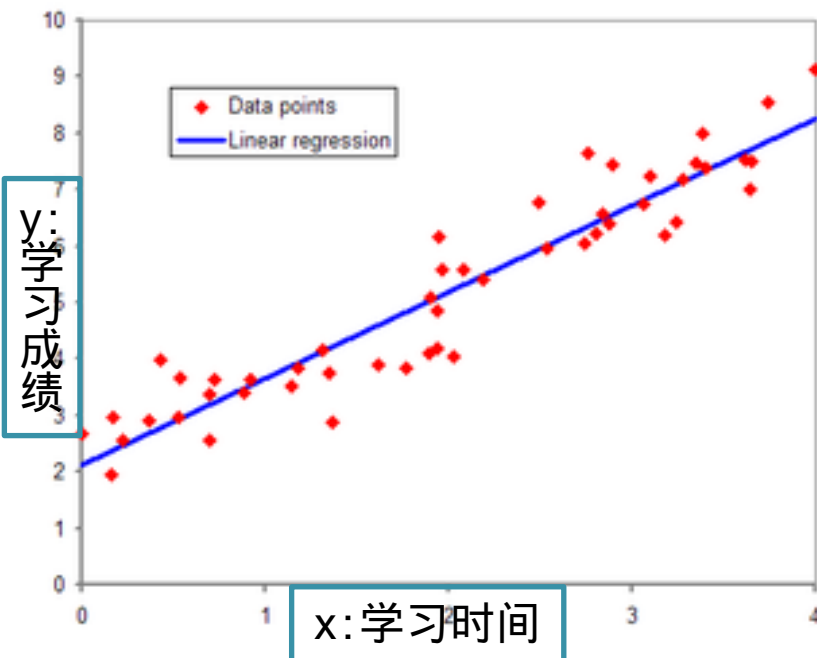
( , 狗)( , 狗) ( , 猫)( , 猫 )  
( , 狗)( , 狗) ( , 猫)( , 猫 )

$(x_1, y_1), (x_2, y_2), \square, (x_n, y_n)$



# 怎样学习

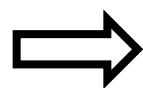
## 线性回归问题



$$\min_{f \in \{ax+b\}} \sum_{i=1}^n (f(x_i) - y_i)^2$$

# 什么是机器学习

要学什么



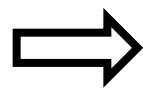
决策函数

从哪里学



训练数据

怎样学习



求解机器学习模型

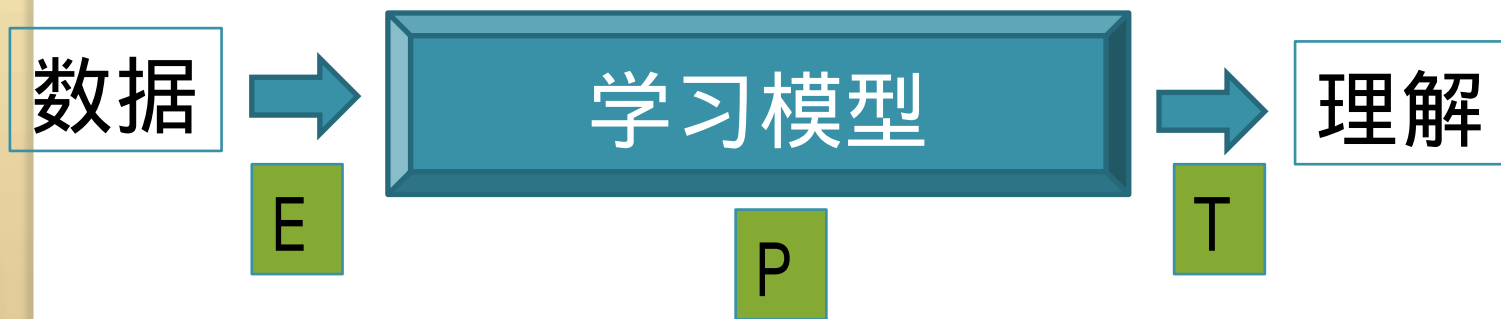


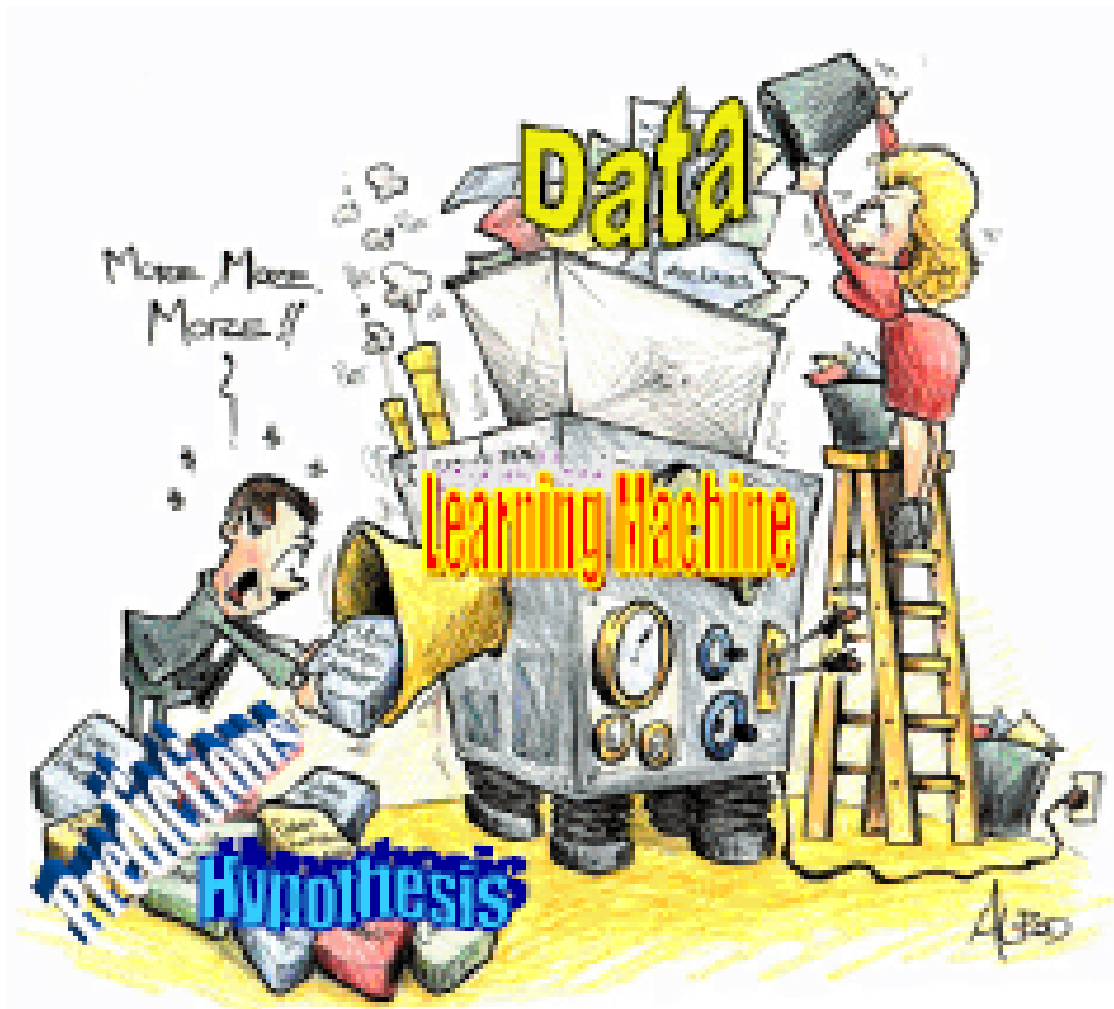
"A computer program is said to learn from **experience E** with respect to some class of **tasks T** and **performance measure P**, if its performance at tasks in T, as measured by P, improves with experience E".

**Tom M. Mitchell, 卡内基梅隆大学, 1997**

## ➤ 机器学习三要素：

- ❑ 使用经验数据 (Experience)
- ❑ 针对某个目标 (Task)
- ❑ 提高学习性能 (Performance)





# 什么是机器学习?

➤ 1959, Arthur Samuel: machine learning is a "Field of study that gives computers the ability to learn without being explicitly programmed".



# 什么是机器学习?

- 1996, Langley: Machine learning is a science of the artificial. The field's main objects of study are artifacts, specifically algorithms that improve their performance with experience.
- 2004, Alpaydin: Machine learning is programming computers to optimize a performance criterion using example data or past experience.

# 什么是机器学习?

Machine Reading



Auto-Text to Knowledge

T: 预测

P:  
学习

E 数据





# 什么是机器学习?

## ➤ 核心

➤ 已知  未知

➤ 有限  无限

# 什么是机器学习?

## ➤ 核心

➤ 已知  未知

➤ 有限  无限

## ➤ 与人学习原理一致

➤ E: 见识多

➤ T: 目标明确

➤ P: 学习方法好

# 什么是机器学习?

**Generalization!**

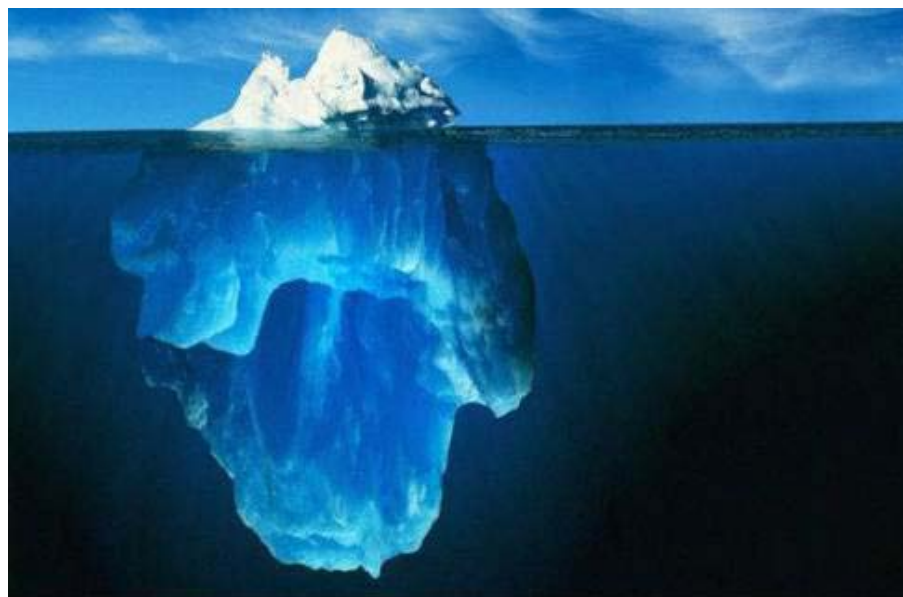


➤ 推广/预测/泛化 能力

# 什么是机器学习?

**Generalization**

**Understanding**



➤ 机器学习角度

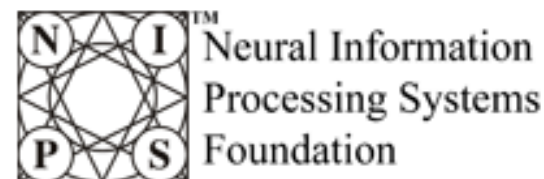
➤ 统计角度

# 机器学习历史

- 1950-1960 兴起
- 1960-1980 沉寂
- 1980-2010 复兴
  - ICML 1980创办
  - NIPS 1987创办, 2000转向ML
  - ML 1986创办
  - JMLR 2000创办
- 2010-今 火热
  - 深层网络
  - 大数据
  - 人工智能种种猜测



International Conference on Machine Learning



**JMLR**

**Journal of Machine Learning Research**

[Home Page](#)

[Papers](#)

[Submissions](#)

[News](#)

[Editorial Board](#)

[Announcements](#)

[Proceedings](#)

[Open Source Software](#)

[Search](#)

[Statistics](#)

The Journal of Machine Learning Research (JMLR) provides an international forum for the electronic and paper publication of high-quality scholarly articles in all areas of machine learning. All published papers are freely available online.

JMLR has a commitment to rigorous yet rapid reviewing. Final versions are published electronically (ISSN 1533-7928) immediately upon receipt. Until the end of 2004, paper volumes (ISSN 1532-4435) were published 8 times annually and sold to libraries and individuals by the MIT Press. Paper volumes (ISSN 1532-4435) are now published and sold by Microtone Publishing.

## News

- Volume 15 completed; Volume 16 began.



# 机器学习典型应用

## ➤ 垃圾邮件过滤

西安交通大学 XIAN JIAOTONG UNIVERSITY

Deyu Meng <dymeng@xjtu.edu.cn> 自助查询 - 锁屏

设置 ▾ | 帮助 | 退

欢迎页 垃圾邮件 × 数学与统计... ×

收信 写信

垃圾邮件 (共 32 封 32 封 未读邮件, 全部设为已读)

	发件人	主题
今天 (20 封)		
<input type="checkbox"/>	choukema	SCI-应用数学和计算_dymeng7282
<input type="checkbox"/>	IEEE Spectru...	Download these 6 white papers courtesy of Masterbond
<input type="checkbox"/>	Edmund	An authoritative forum will be held in Guilin on April 21-22, make the contribution before 3.1
<input type="checkbox"/>	ICPCS Conf.	通知
<input type="checkbox"/>	aamservice	《国际应用数学进展》编委邀请函孟德宇老师
<input type="checkbox"/>	Springer Nat...	Deyu Complete our book author survey with a chance to win
<input type="checkbox"/>	Amanda Venis	Publish your medical research data
<input type="checkbox"/>	ICPCS Conf.	通知
<input type="checkbox"/>	Frontiers of C...	Reminder: review invitation Frontiers of Computer Science
<input type="checkbox"/>	J submit	欢迎您成为编委/审稿人。亲爱的Author, Robust low-rank tensor factorization by cyclic weighted median
<input type="checkbox"/>	ResearchGate	Deyu, you were recently cited by an author from University of Western Australia

系统维护: 西安交通大学网络中心 服务电话 82667777 邮箱 support@xjtu.edu.cn

# 机器学习典型应用

## 检索

Google deyu meng

全部 新闻 视频 图片 更多 设置 工具

找到约 71 条结果 (用时 0.70 秒)

**西安交通大学教师个人主页 - Deyu Meng's Homepage (孟德宇的主页)**  
dymeng.gr.xjtu.edu.cn/

**Deyu Meng (孟德宇)** 教授, 博导, Professor 西安交通大学 大数据算法与分析技术国家工程实验室 西安数学与数学技术研究院 西安, 陕西, 710049. National ...

**Deyu Meng - Google Scholar Citations**  
scholar.google.com/citations?user=an6w-64AAAAJ&hl=zh-CN 翻译此页  
L Jiang, D Meng, Si Yu, Z Lan, S Shan, A Hauptmann. Advances in Neural Information Processing Systems, 2078-2086, 2014. 149, 2014. A generalized iterated ...  
您访问过该网页很多次。上次访问日期: 16-2-18

**Deyu Meng's Homepage, Meng Deyu's Homepage,**  
www.cs.cmu.edu/~deyum/ 翻译此页  
**Deyu Meng (孟德宇)** Email: dymeng AT mail Dot xjtu DOT edu DOT cn. Address: Institute for Information and System Sciences School of Mathematics and ...

**dblp: Deyu Meng**  
https://dblp.uni-trier.de/ Persons  
List of computer science publications by **Deyu Meng**.

**Deyu Meng | Xi'an Jiaotong University, Xi'an | XJTU - ResearchGate**  
https://www.researchgate.net/ Xi'an Jiaotong University  
**Deyu Meng** of Xi'an Jiaotong University, Xi'an (XJTU). Read 95 publications, and

国内版 国际版

baidu deyu meng

网页 图片 视频 学术 词典 地图

14,800 条结果 时间不限

**Deyu Meng's Homepage, Meng Deyu's Homepage,** 翻译此页  
2019-2-16 · Qian Zhao, Deyu Meng, Zongben Xu, Wangmeng Zuo, Lei Zhang. Robust principal component analysis with complex noise, Supplementary Material, ICML, 2014. Matlab code. Our proposed FastMMD method decreases the time complexity of MMD calculation from conventional  $O(N^2d)$  to  $O(N \log(d))$ . Ji Zhao, Deyu Meng.  
www.cs.cmu.edu/~deyum

**西安交通大学教师个人主页 - My Research (我的研究)**  
New: [118] Jing Yao, Deyu Meng, Qian Zhao, Wenfei Cao, and Zongben Xu. Nonconvex-sparsity and Nonlocal-smoothness Based Blind Hyperspectral Unmixing. IEEE ...  
www.gr.xjtu.edu.cn/web/dymeng/2

**西安交通大学教师个人主页 - Project Page: Self-paced ...**  
This paper provides some fundamental understanding of various SPL regimes. Deyu Meng, Qian Zhao, Lu Jiang. A Theoretical Understanding of Self-paced Learning. accepted ...  
www.gr.xjtu.edu.cn/web/dymeng/6

**Deyu meng - Lecturer - Xi'an Jiaotong University | 领英**  
职位: Associate Professor at Xi'an ... 行业: 教育系统  
人数: 124 位置: 中国 陕西 西安  
上领英, 在全球领先职业社交平台查看 Deyu meng 的职业档案。Deyu 的职业档案列出了 2 个职位。查看 Deyu 的完整档案, 结识职场人脉和查看相似公司的职位。  
https://cn.linkedin.com/in/deyu-meng-4921b120

**dblp: Deyu Meng** 翻译此页  
Xiai Chen, Zhi Han, Yao Wang, Qian Zhao, Deyu Meng, Lin Lin, Yandong Tang: A Generalized Model for Robust Tensor Factorization With Noise Modeling by Mixture of

Baidu deyu meng

网页 资讯 贴吧 知道 视频 音乐 图片 地图 文库 更多 搜索工具

检索为您找到相关结果约 27,500 个

**西安交通大学教师个人主页 - Deyu Meng's Homepage (孟德宇的主页)**

 2019年1月8日 - 西安交通大学教师个人主页平台是国内第一家实时发布本校教师信息的官方网站, 由西安交通大学主办, 为广大用户第...  
gr.xjtu.edu.cn/web...meng - 百度快照

**Deyu Meng's Homepage, Meng Deyu's Homepage,**  
查看此网页的中文翻译, 请点击 翻译此页  
**Deyu Meng (孟德宇)** Email: dymeng AT mail Dot xjtu DOT edu DOT cn Address: Institute for Information and System Sciences School of Mathematics and ...  
www.cs.cmu.edu/~deyum/... - 百度快照

**抑郁 百度百科**

 这是负性情感增强的表现, 患者自觉情绪低沉, 整日忧心忡忡, 对自我才智能力估计过低, 对周围困难估计过高。本症患者常伴有思维迟缓, 言语动作减少, 意志活动减退。这一症状为躁...  
**病因 临床表现 辅助检查 治疗 预防 研究发现**  
baike.baidu.com/

**心烦、郁闷的时候, 怎么办? 百度经验**





# 机器学习典型应用

## ➤ 推荐系统

[新闻](#) [网页](#) [贴吧](#) [知道](#) [音乐](#) [图片](#) [视频](#) [地图](#) [百科](#) [文库](#) [更多>>](#)

百度一下

[劳动节不劳动，吃喝玩乐5.1元起!](#)

导航

新闻

音乐

视频



[猜你喜欢](#) [短视频](#) [综艺](#) [电视剧](#) [电影](#) [动漫](#) [游戏直播](#)



- 1 黄子韬与SM公司解约 其父发文证实 1222679
- 2 实拍男子骑自行车飞跃兰博基尼 车... 592827
- 3 爱奇艺早班机：EXO黄子韬解约 独... 525112
- 4 北京台批《名侦探柯南》：暴力肆虐... 285495
- 5 中国首次发布红色通缉令 缉拿百名... 246957
- 6 娱乐猛回头：Baby吃定晓明有心机 ... 141790
- 7 王诗龄嘟嘴卖萌 被司机扛起开心玩耍 71404
- 8 两岁娃被送错校车 在新幼儿园他淡... 62602
- 9 大S救火代班《康熙》 才录第一集... 51566



# 机器学习典型应用

## ➤ 用户甄别



# 机器学习典型应用

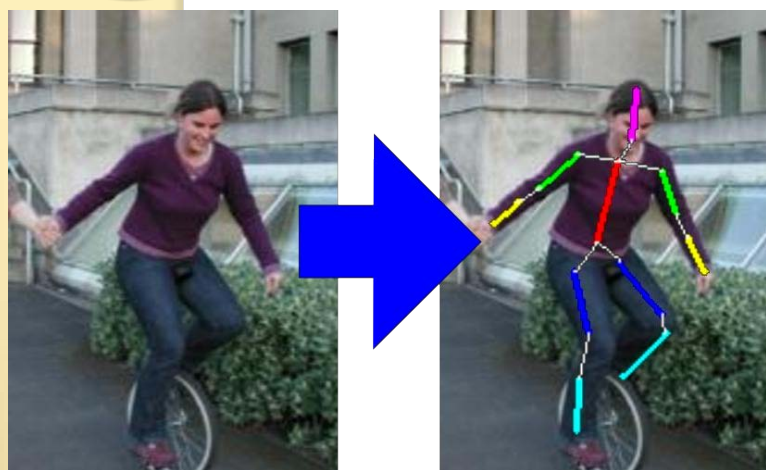
## ➤ 物体识别



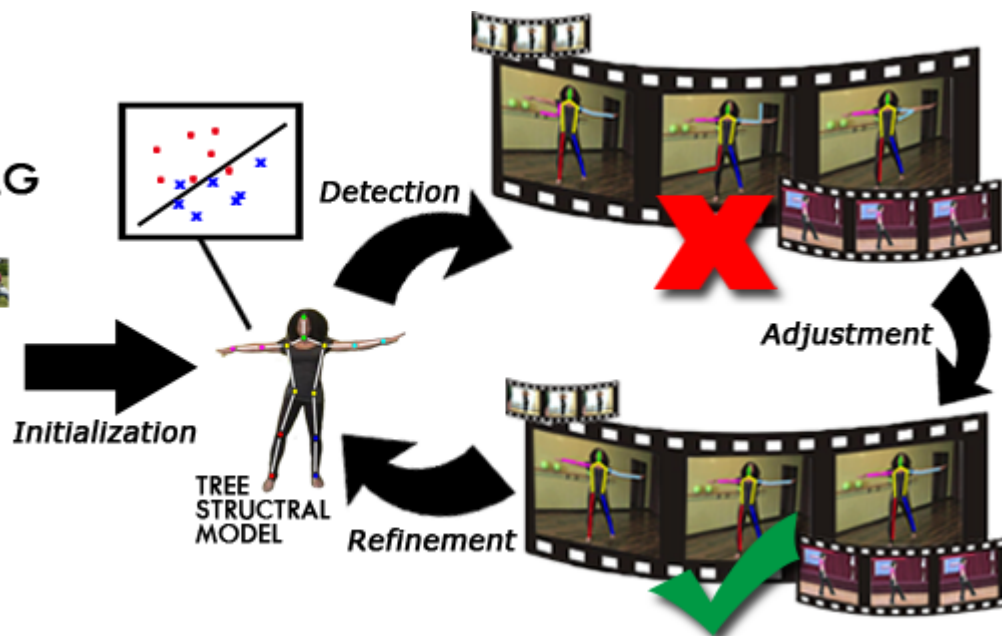


# 机器学习典型应用

## ➤ 姿态识别

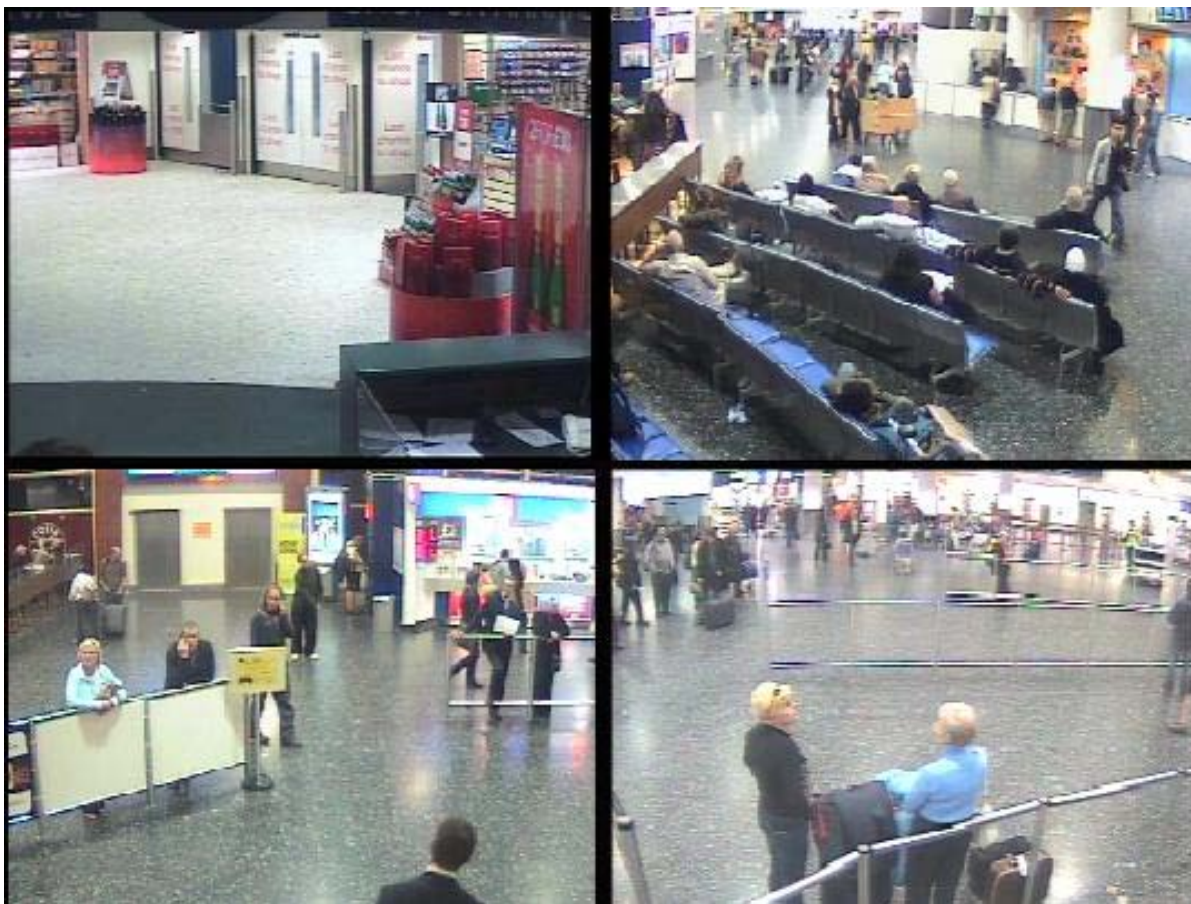


LABELLED IMG



# 机器学习典型应用

## ➤ 事件检测



# 典型机器学习问题

## ➤ 机器学习问题的基本要素

### ➤ 训练数据集

$$\{x_i, y_i\}_{i=1}^n$$

### ➤ 未知的目标函数

$$f: X \rightarrow Y$$

### ➤ 学习机器(备选函数集)

$$H = \{h/h: X \rightarrow Y\}$$

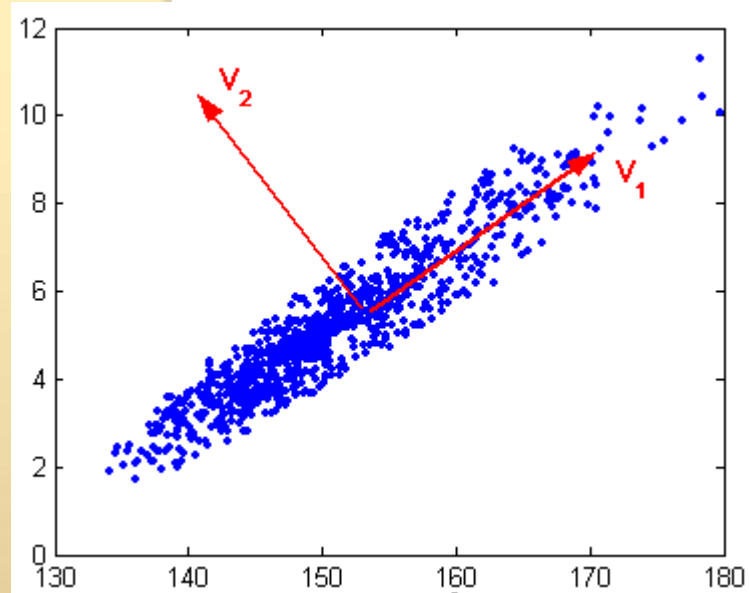
## ➤ 机器学习问题设置:

### ➤ 输入: 训练数据与预设学习机器

### ➤ 输出: 学习机中最接近未知目标函数的逼近

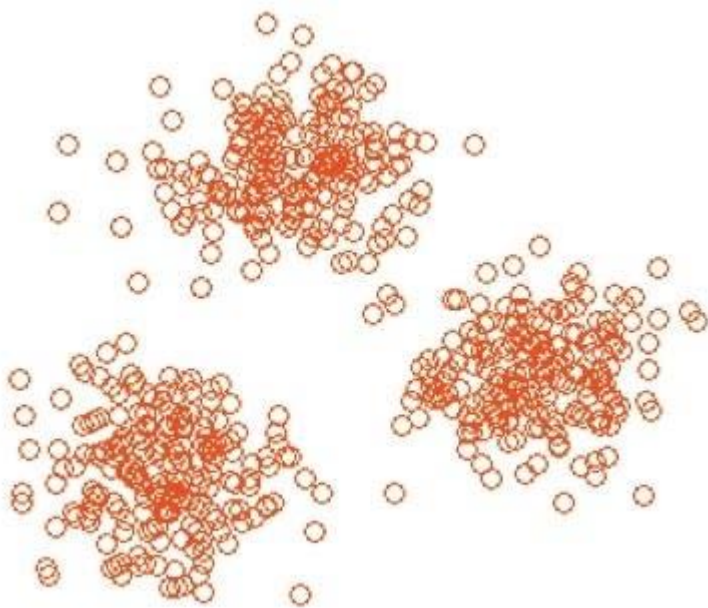
# 典型机器学习问题

## ➤ 降维



# 典型机器学习问题

## ➤ 聚类

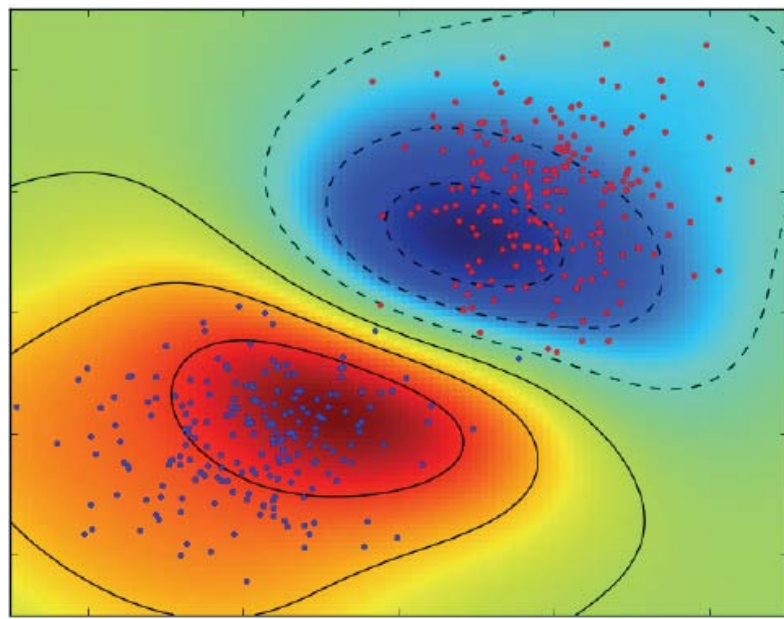
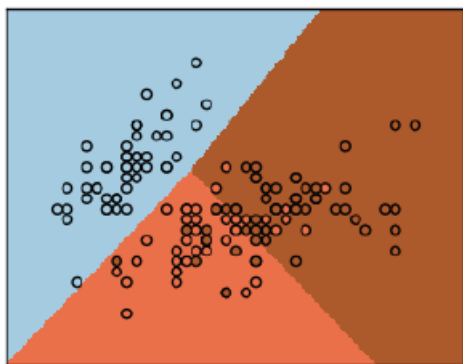
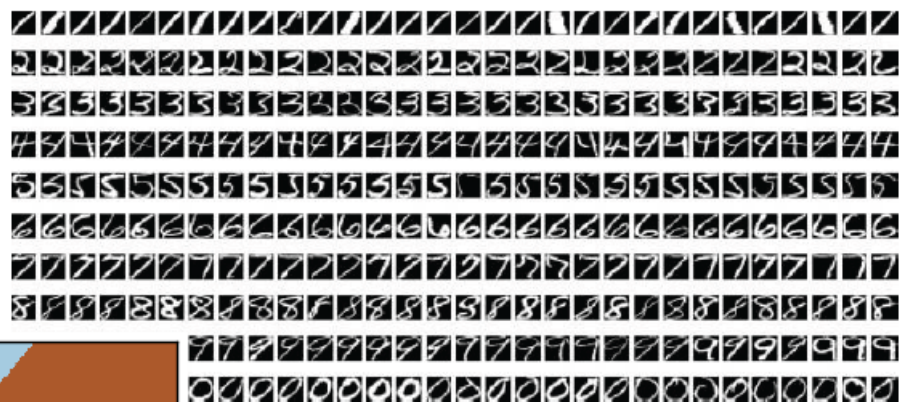


- 文本
- 用户
- 网页
- 疾病
- 图片
- 交通工具
- ○ ○ ○ ○ ○ ○ ○



# 典型机器学习问题

## ➤ 分类

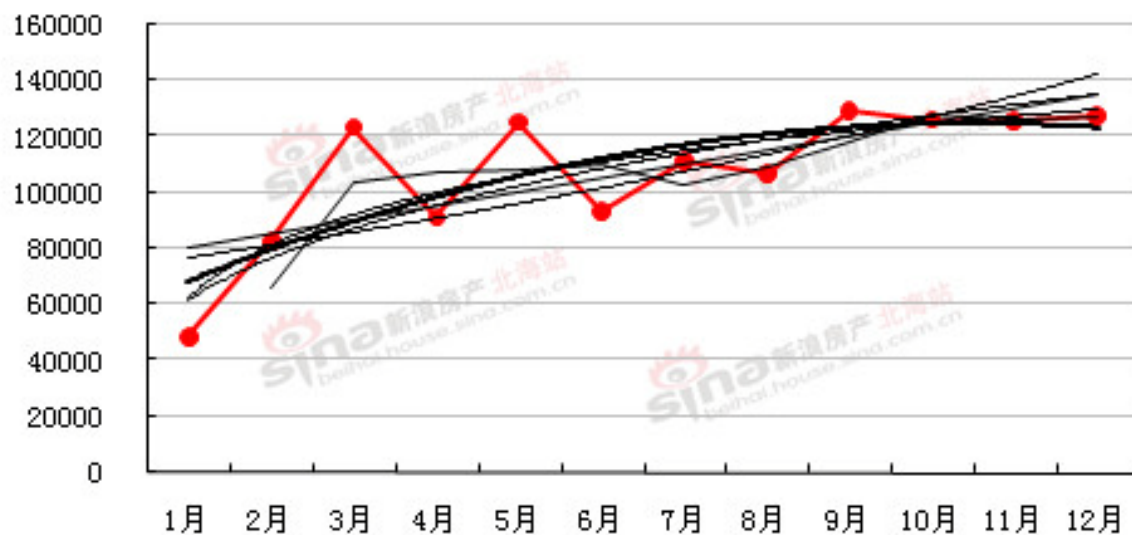




# 典型机器学习问题

## ➤ 回归

2012年北海网签面积趋势预测



# 本课程计划讲授内容

## (根据教学情况适当缩减)

介绍:

I-1. 介绍

I-2. 决策树

I-3. 统计基础: MLE + MAP

聚类:

I-4. k-均值

I-5. 编程展示

I-6. 扩展

# 本课程计划讲授内容

降维:

I-7.PCA + SVD

I-8. 编程演示

I-9. 扩展

有监督学习: (赵谦老师)

I-10. 支撑向量机

I-11. 编程演示

I-12. 扩展

学术前沿: (if possible)

I-13. 动态



Enjoy!

# 要求

1. 搞清楚机器学习的基本定义
2. 搞清楚机器学习的典型问题
3. 搞清楚学习的核心要素

## 阅读：

- [1] Pattern Recognition and Machine Learning, Christopher , M. Bishop, Springer, 2006. I. Introduction
- [2] High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality, David L. Donoho, ICM 2000
- [3] 周志华，机器学习，清华大学出版社，2015.