

机器学习与数据科学

Machine Learning and Data Science

主讲: 李春光

www.pris.net.cn/teacher/lichunguang

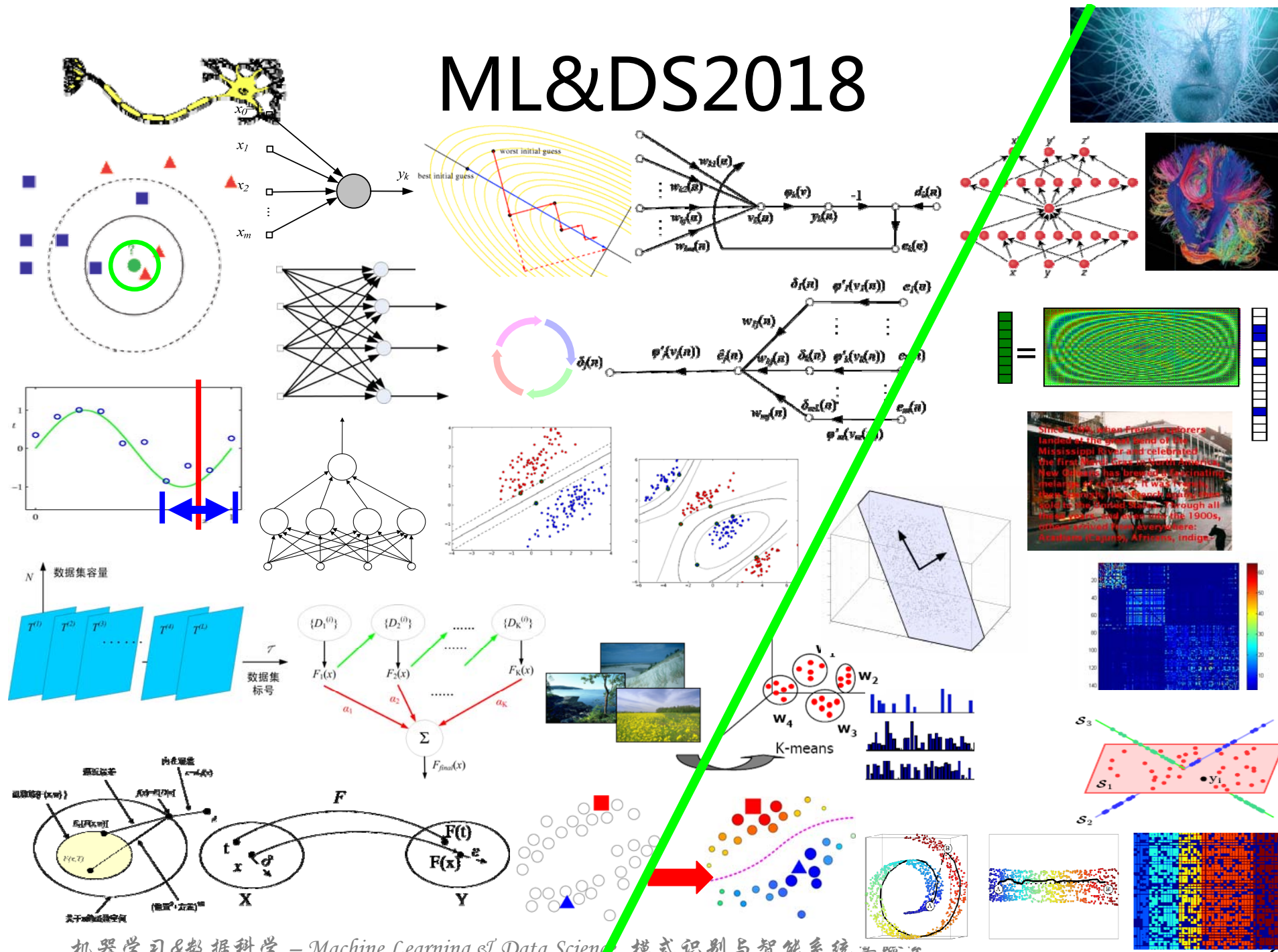
模式识别与智能系统实验室

信息与通信工程学院 网络搜索教研中心

北京邮电大学



ML&DS2018



内容提要

- 引言
- 机器学习的发展简史
- 机器学习相关领域国际会议和期刊
- 机器学习相关参考资料

引子: 从学习说起...

- 思考一下我们日常生活中的学习过程

- 什么是学习

- 子曰: “学而时习之, 不亦说乎” ?
 - “学”: 闻、见, 是获取信息的过程; “习”: 温习、练习、实践
 - 学习: 获取知识、形成技能, 获得适应环境、改变环境的能力的过程

- 学习是怎样一个过程

- 怎样评价学习的效果

- 怎样有效地学习

引子: 从学习说起...

- 思考一下我们日常生活中的学习过程
 - **什么是学习**
 - 学习: 获取知识、形成技能, 获得适应环境、改变环境的能力的过程
 - **学习是怎样一个过程**
 - 通过了解、认识、记忆、理解 “问题与答案” 来进行, 进而达到能够自由运用与推广
 - **怎样评价学习的效果**
 - **怎样有效地学习**

引子: 从学习说起...

- 思考一下我们日常生活中的学习过程

- **什么是学习**

- 学习: 获取知识、形成技能, 获得适应环境、改变环境的能力的过程

- **学习是怎样一个过程**

- 通过了解、认识、记忆、理解 “问题与答案” 来进行, 进而达到能够自由运用与推广

- **怎样评价学习的效果**

- 测试

- 测试题目是原封不动所学过的: 考察记忆能力
 - 测试题目是学过的但改变了形式的: 考察运用与推广能力

- **怎样有效地学习**

引子: 从学习说起...

- 思考一下我们日常生活中的学习过程
 - **什么是学习**
 - 学习: 获取知识、形成技能, 获得适应环境、改变环境的能力的过程
 - **学习是怎样一个过程**
 - 通过了解、认识、记忆、理解 “问题与答案” 来进行, 进而达到能够自由运用与推广
 - **怎样评价学习的效果**
 - 测试
 - **怎样有效地学习**
 - 先测试、后学习
 - 当我们绞尽脑汁地不断思考——“它是怎么回事”之后, 才能够进入**高效学习**的状态。

引子: 从学习说起...

- 思考一下我们日常生活中的学习过程
 - **什么是学习**
 - 学习: 获取知识、形成技能, 获得适应环境、改变环境的能力的过程
 - **学习是怎样一个过程**
 - **怎样评价学习的效果**
 - **怎样有效地学习**
- 如何让计算机像人类一样具有学习能力?
 - **人工智能**
 - 应用层面: 模式识别, 机器视觉, 数据挖掘, 信息检索...
 - 理论层面: 机器学习, 数据科学

什么是机器学习

- 机器学习的几种定义:
 - “机器学习是一门人工智能的学科，主要研究如何通过经验改善具体算法的性能”
 - 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.**
 - “机器学习是对能通过经验自动改进的计算机算法的研究”
 - Machine Learning is the study of computer algorithms **that improve automatically through experience.**
 - Tom Mitchell, 1997
 - “机器学习是用数据或以往的经验，去优化计算机程序的性能准则”
 - Machine learning is programming computers **to optimize a performance criterion using example data or past experience.**

Machine learning

A computer program is said to learn from experience E

with respect to some tasks T and

performance measure P ,

if its performance at tasks in T ,

as measured by P ,

improves with experience E .

机器的学习过程

- 从经验观测数据中获得事物内在模式或规律以自我改进的过程

A computer program is said to learn from experience E

**with respect to some tasks T and
performance measure P,
if its performance at tasks in T,
as measured by P,
improves with experience E.**

什么是数据科学

- TBD: To be defined
 - 2015年秋, Dr. Afonso S. Bandeira 在MIT开设了一门课程——Topics in Mathematics of Data Science
 - 面向对theoretical aspects of algorithms that aim to extract information from data方向做研究感兴趣者
 - 内容覆盖：
 - Dimension reduction / Manifold learning / Semi-supervised Learning
 - Concentration Inequalities / Sparse recovery / Random Matrices / Approximation Algorithms
 - Community detection in graphs, and several others

什么是数据科学

- TBD: To be defined
 - 2015年秋, Dr. Afonso S. Bandeira 在MIT开设了数据科学在狭义上对应于机器学习的无监督学习部分；还包括数据的感知与获取、(探索性)数据分析、数据表示与处理、模式发现等
 - 内容覆盖：
 - Dimension reduction / Manifold learning / Semi-supervised Learning
 - Concentration Inequalities / Sparse recovery / Random Matrices / Approximation Algorithms
 - Community detection in graphs, and several others

机器学习与数据科学

- 从数据中学习 (Learning from Data)
 - 从经验观测数据中获得事物内在模式或规律

请思考两个问题

- 为什么要让计算机从数据中学习？
 - 回答: ...
 - **有些工作需要计算机代替或帮助人来完成**
 - 光学字符识别(OCR)、监控、医学诊断、信用审查
 - Web搜索...
 - 个性化推荐...
- 这类问题有哪些共同特点？
 - **能够收集到观测数据**
 - **数据中存在模式或规律**
 - **无法直接写出公式**
 - why ?

内容提要

- 引言
- 机器学习的发展简史
- 机器学习相关领域国际会议和期刊
- 机器学习相关参考资料

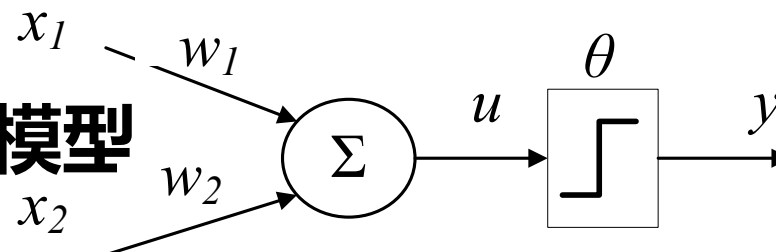
机器学习的发展简史

- 初始发展期(40年代-50年代)
- 感知器时代:(60年代前后)
 - **神经网络的第一次研究热潮**
- 感知器的衰落与学习理论基础的奠定(60年代末-80年代初)
 - **感知器局限性被发现**
- 多层感知器的复兴: 神经网络时代(80年代以后)
 - **神经网络的第二次研究热潮: 多层感知器**
- 统计学习理论完善与成熟: (90年代中期以后)
 - **正则化框架, 统计学习理论和SVM横空出世**
- 无监督学习和半监督的研究热潮(Since 2000)
 - **流形学习, 压缩感知**
- 深度学习: 神经网络卷土重来(since 2006)

初始发展期：M-P模型诞生

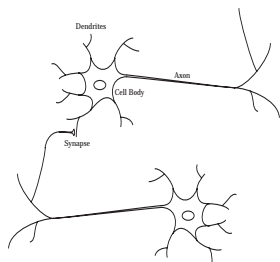
- 1943年，McCulloch和Pitts的开创性工作

– 提出兴奋与抑制型神经元模型



- 对输入信号加权求和，再与阈值比较以确定输出
- 结合神经生理学和数理逻辑，描述了神经网络的逻辑演算，并证明这样的网络原则上可计算任何数学和逻辑函数

- McCulloch：精神病学家，解剖学家
- Pitts: 数学天才



[1] McCulloch W.S. and Pitts W., "A logical calculus of the ideas immanent in nervous activity" . Bulletin of Mathematical Biophysics, vol.5, 1943, pp.115-133.

– 标志着神经网络和人工智能学科的诞生

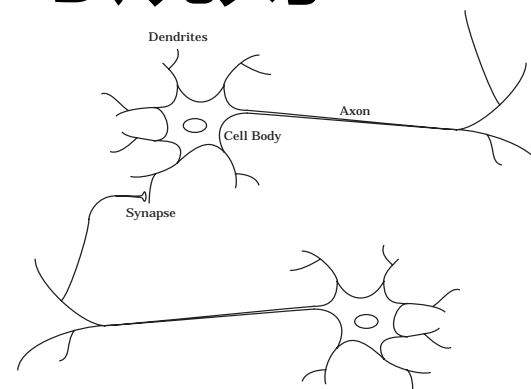


初始发展期：Hebb学习规则

- 1949年 Hebb的开创性工作

- **发表《行为组织学》**

- The Organization of Behavior



- **第一次清楚的说明了突触修正的生理学学习规则**

- 认为信息存储在神经元的连接中，并提出一个调节连接权值的学习策略
 - 著名的学习假说：
 - “当细胞A的轴突到细胞B的距离近到足够激励它，且反复或持续地刺激B，那么在这两个细胞或一个细胞中会发生某种增长过程或代谢作用，增加A对细胞B的刺激效果”
 - 两个神经元之间的可变突触的作用被突触两端神经元中一个对另一个的重复激活所加强

神经网络的第一次研究热潮时期

- 1958年 Rosenblatt : 感知器(Perceptron)的提出
 - **第一次定义面向计算的神经网络**
 - **引入了能够训练神经网络解决模式分类问题的学习规则**
 - 并证明：只要求解问题的权值存在，通常会收敛到正确的网络权值上

[1] Rosenblatt F. The perceptron: probabilistic model for information storage and organization in the brain [J]. Psychological Review, 1958, 65(6):386-408

[2] Albert B.J. **Novikoff**. On convergence proofs on perceptrons. In Proc. of Symposium on the Mathematical Theory of Automata, 12, 615–622, 1962.

- 1960年 Widrow and Hoff : 自适应线性单元(Adaline: adaptive linear element)
 - **使用最小均方(LMS: Least Mean Square)规则完成训练**
 - **与感知器极为相似，可用于自适应控制和模式识别**

[3] B. Widrow and M. E. Hoff. Adaptive switching circuits. In IRE WESCON Convention Record, volume 4, pages 96--104, 1960.

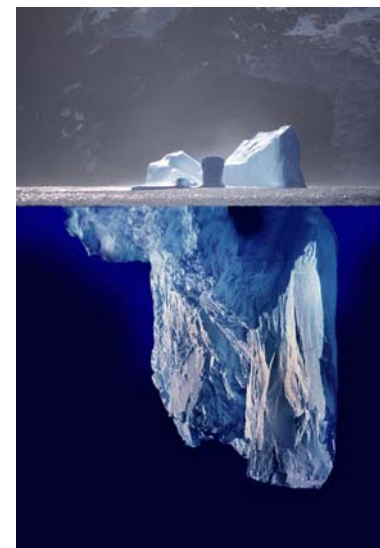
感知器研究的低谷

- 1969年 Minsky and Papert
 - 发表了《感知器(Perceptrons)》一书
 - 质疑单层神经网络能力有限，认为感知器的缺陷无法克服
 - 导致神经网络研究的第一次低潮
- 1972年Kohonen和Anderson
 - 分别独立提出按内容可寻址联想记忆模型
- 1974年 Werbos
 - 给出训练多层前馈感知器的反向传播(BP: back propagation)算法的首次描述



Parker, 1985, TR; LeCun, 1985, Cognitiva (French)

Rumelhart, Hinton & Williams, 1986



学习理论基础的奠定(60-70年代)

- 结构风险最小化原则的理论

- 1968年，**VC维**的概念和**泛函空间的大数定律**

- Vapnik & Chervonenkis 1968.

- 1974年，结构风险最小化归纳原则

- Vapnik V. and Chervonenkis A. Theory of Pattern Recognition (In Russian), Nauka, Moscow, 1974.



- 正则化理论

- 在解决一个**不适定问题**时，通过最小化一个**正则化泛函**，会获得一个收敛于真实解的解序列

- Tikhonov: 1963

- Ivanow: 1962

- Phillips: 1962.

使用正则化泛函作为目标函数

神经网络研究的第二次研究热潮

- 1985年Ackley, Hinton和Sejnowski提出Boltzmann机
- 1985年Parker和LeCun分别独立提出反向传播算法
 - Parker 1985, Learning-Logic, Technical Report;
 - LeCun1985: Cognitiva, in French
- 1986年Rumelhart, Hinton和Williams再次提出反向传播算法, 论文发表于Nature
 - 1987年Sivilotti, Mahowald和Mead 神经网络的第一次VLSI实现
 - 1988年 Broomhead和Lowe第一次把径向基函数引入神经网络
- 1982年 Hopfield 开启神经网络现代史
 - 描述了递归神经网络的运行方式和功能
 - Hopfield 联想记忆网络:
 - 给定一个有噪声的、不完整的模式, 网络能够输出所记忆的完整模式
- 1982年Kohonen: 自组织特征映射(SOM)
 - 1982年Oja利用Hebb规则训练单个线性神经元提取主成分
 - 1985年Herault, Jutten和Ans把神经网络用于独立源信号分离



Geoffrey E. Hinton

Rumelhart, Hinton & Williams, “Learning representations by back-propagating errors”, Nature, 1986

正则化理论

- 20世纪90年代初
 - Tomaso Poggio 等把正则化理论引入到学习问题

$$H[f] = \sum_{i=1}^N (d_i - f(\xi_i))^2 + \lambda \|Pf\|^2 \quad (1)$$



- [1] Poggio, T. and Girosi, F. (1990): "Regularization Algorithms for Learning that are Equivalent to Multilayer Networks," [Science](#), Vol. 247, No. 4945. Feb. 23, 1990, pp. 978-982.
- [2] Wahba, G. 1990: Spline Models for Observational Data, SIAM, Philadelphia.
- [3] Girosi, F., M. Jones, and T. Poggio. Regularization Theory and Neural Network Architectures." *Neural Computation* 7 (1995): 219-269
- [4] Poggio, T. & Smale, S. The mathematics of learning: Dealing with data. *Nat. Am. Math. Soc.* 50, 537–544 (2003).
- [5] Cucker, F. & Smale, S. On the mathematical foundations of learning. *Bull. Am. Math. Soc.* 39, 1–49, (2001).

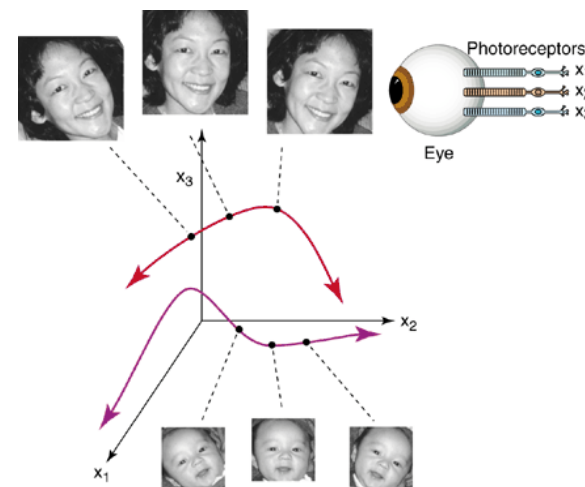
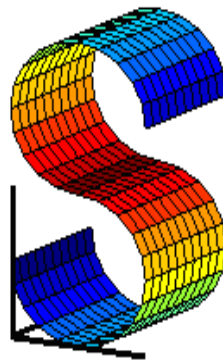
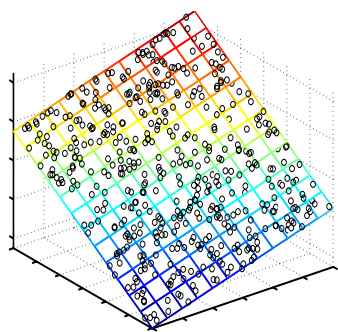
统计学习理论的成熟与SVM横空出世

- 20世纪90年代早期
 - Vapnik和他的合作者发明了支持向量机，把研究热点引向了统计学习理论和机器学习理论
 - **SVM: Support Vector Machine**
 - Boser, Guyon and Vapnik, “[A training algorithm for optimal margin classifiers](#)”, 5th Annual Workshop on Computational Learning Theory, pp.144-152, 1992.
 - Cortes and Vapnik, “[Support-Vector Networks](#)”, Machine Learning, Vol.20, pp.273-297, 1995.
 - **SLT**
 - V. Vapnik, The nature of statistical learning theory, New York, Springer-Verlag, 1995.
 - V. Vapnik, Statistical learning theory, New York, Wiley, 1998.



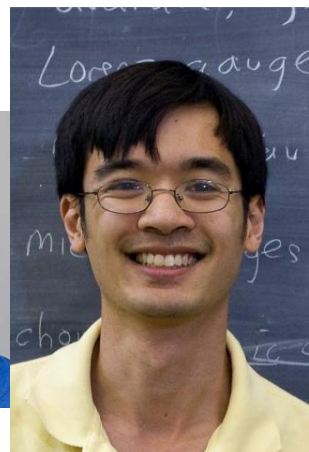
无监督学习的兴起

- 2000年：流形学习(Manifold Learning)
 - 2000年12月，Science杂志上3篇流形学习论文的连载，揭开了流形学习研究热潮的序幕
 - Roweis et al. SCIENCE 2000;
 - Tenenbaum et al. SCIENCE 2000;
 - Seung & Lee, SCIENCE 2000;



压缩感知的理论突破

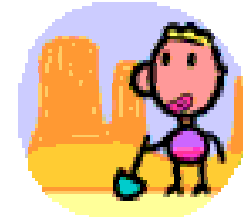
- 2006年：压缩感知(Compressive Sensing)
 - 稀疏编码(Sparse Coding):
 - 2006年，E. Candes & T. Tao / Donoho 等人证明了在**一定条件下**，**L1最小化**问题的解是原**L0问题**的最优解
 - 矩阵填充(Matrix Completion):
 - 2009年，E. Candes & B. Recht 证明在**一定条件下**，通过求解**核范数**最小化的**凸优化**问题能够从**一部分观测**中**准确获得**该矩阵。
 - Robust PCA?



神经网络卷土重来的前奏

New Life for Neural Networks

Garrison W. Cottrell



Recent advances in machine learning have caused some to consider neural networks obsolete, even dead. **This work** suggests that such announcements are premature.

选读：ANN死了么？No.

[1] Garrison W. Cottrell, New Life for Neural Networks, SCIENCE, Vol.313,

[2] **Hinton G.E. and Salakhutdinov R.R., Reducing the dimensionality of data with neural networks, SCIENCE, vol.313, July 2006, pp.504-507.**

神经网络卷土重来: 第3次研究热潮



- 深度学习 (1.0)
 - 深度置信网络(DBNs)
 - 2006 : Hinton' s revolutionary work on Deep Belief Networks (DBNs)
 - 自编解码网络
 - 2006 : encoder-decoder
 - 稀疏自编解码网络
 - 2006 : Sparse encoder-decoder
- 三篇点火之作:



[1] Hinton et al., “A fast learning algorithm for deep belief nets”, Neural Computation, 18:1527-1554, 2006.

[2] Yoshua Bengio : “Greedy Layer-Wise Training of Deep Networks”, NIPS 2006.

[3] Yann LeCun: “Efficient Learning of Sparse Representations with an Energy-Based Model”, NIPS 2006.

- Yoshua Bengio, “Learning Deep Architectures for AI” , Foundations and Trends in Machine Learning, 2 (1), 2009.

机器学习的发展简史

- 初始发展期(40年代-50年代)
- 感知器时代:
 - **神经网络的第一次研究热潮(60年代前后)**
- 感知器的衰落与学习理论基础的奠定(60年代末-80年代初)
 - **感知器局限性被发现**
- 多层感知器的复兴: 神经网络时代(80年代以后)
 - **神经网络的第二次研究热潮: 多层感知器**
- 统计学习理论完善与成熟: (90年代中期以后)
 - **正则化框架, SVM横空出世**
- 无监督学习和压缩感知的研究热潮(Since 2000)
 - **流形学习, 压缩感知**
- 深度学习: 神经网络卷土重来(since 2006 / 2009 /2012)

内容提要

- 引言
- 机器学习的发展简史
- 机器学习相关领域国际会议和期刊
- 机器学习相关参考资料

机器学习主要国际会议

- ICML / NIPS / COLT
 - 机器学习领域三大会议
 - COLT: Conf. On. Learning Theory
 - ICML: International Conf. on Machine Learning
 - NIPS: Annual Conference in Neural Information Processing System
- ECML / UAI / AISTATS / (ACML)
 - ECML: European Conf. on Machine Learning
 - UAI: Intern. Conf. on Uncertainty in Artificial Intelligence
 - AISTATS: Intern. Conf. on Artificial Intelligence and Statistics
 - UAI和AI&STATS侧重于概率推理
 - ACML: Asian Conference on Machine Learning

机器学习主要国际期刊

- JMLR:
 - Journal of Machine Learning Research
- IEEE TIT:
 - IEEE Transaction on Information Theory
- IEEE TPAMI:
 - IEEE Transaction on Pattern Analysis and Machine Intelligence
- 其它:
 - Machine Learning
 - Neural Computation
 - IEEE Trans. Neural Networks & Learning System
 - Neural Networks / Neurocomputing

[修正版]人工智能领域会议介绍-1

- **COLT (1):**
 - ACM主办, 每年举行. 理论计算机科学和机器学习的交叉
- **ICML (1):**
 - 机器学习方面最好的会议之一. 现在是IMLS主办, 每年举行
- **NIPS (1):**
 - NIPS主办, 每年举行. 举办地几乎是一样的
- **UAI (1-):**
 - 表示\推理\学习, AUAI (Association of UAI) 主办, 每年开
- **AISTATS(1-) :**
 - 人工智能与统计, 每年召开
- **ECML (1-):**
 - 机器学习方面仅次于ICML的会议
- **IJCAI (1-):**
 - AI最好的综合性会议, 1969年始, 奇数年开, 从2016年开始每年开
- **AAAI (1-):**
 - 美国人工智能学会AAAI的年会. 但其档次不稳定, 可以给到1-或者2+
- **ACML(1-):**
 - 亚洲机器学习大会, 2009年起创办, 每年举行
- **ICLR(1-):** 表示学习国际大会, 2013年创立

[修正版]人工智能领域会议介绍-2

- **ICCV (1):**
 - 计算机视觉方面最好的会议之一，偏机器视觉. IEEE主办. ICCV逢奇数
年开
- **CVPR (1):**
 - 计算机视觉和模式识别方面最好的会议之一，偏模式识别, IEEE主办, 每
年举行
- **ECCV (1-):**
 - 计算机视觉方面仅次于ICCV, 偶数年份开会
- **BMVC(2):**
 - 英国机器视觉大会，每年召开
- **ACCV(2-):**
 - 亚洲的计算机视觉会议, 在亚太级别的会议里算很好的了，偶数年份开会.
 -
- **ACM MM (2-)**
 - ACM多媒体大会

[修正版]人工智能领域会议介绍-3

- **SIGKDD (1-):**
 - 数据挖掘方面最好的会议, ACM主办, 每年开
- **WWW(2+):**
 - 世界互联网大会
- **ICDM (2+):**
 - 数据挖掘方面仅次于SIGKDD的会议.
- **SDM (2+):**
 - 即SIAM DM, 数据挖掘方面仅次于SIGKDD的会议.
- **SIGIR (2+):**
 - 信息检索方面最好的会议, ACM主办
- **PKDD (2-):**
 - 欧洲的数据挖掘会议, 目前在数据挖掘会议里面排第4.
- **ECIR (3+):**
 - 欧洲的信息检索会议, 前几年还只是英国的信息检索会议
- **PAKDD (3+):**
 - 亚太数据挖掘会议, 目前在数据挖掘会议里排第5.

[修正版]人工智能领域会议介绍-4

- **ICASSP (3+):**
 - 语音方面最重要的会议之一
- **ICME (3+):**
 - 多媒体、模式识别方面最著名的盛会之一.
- **ICPR (3+):**
 - 模式识别方面最著名的会议之一, 盛会型.
- **ICIP (3):**
 - 图像处理方面最著名的会议之一, 盛会型.
- **IJCNN (3):**
 - 神经网络方面最重要的会议, 盛会型
- **ICANN (3):**
 - 欧洲的神经网络会议, 从quality来说是神经网络会议中最好的, 但这个领域的人不重视会议, 在该领域它的重要性不如IJCNN.
- **CAI (3):**
 - 加拿大的综合型人工智能会议, 在国家/地区级AI会议中算不错的了.
- **PRICAI (3):**
 - 亚太综合型人工智能会议

内容提要

- 引言
- 机器学习的发展简史
- 机器学习相关领域国际会议和期刊
- 机器学习相关参考资料

神经网络与机器学习书目 (1)

- 神经网络与机器学习
 - **Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, MIT Press, 2016.**
 - **Simon Haykin, Neural Networks and Learning Machines (Third Edition), Pearson Education, 2009.03. [神经网络与学习机器]**
 - **Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.**
 - **Simon Haykin, Neural Networks: a comprehensive foundation (2ed edition), Pearson Education.**
 - Simon Haykin著，叶世伟，史忠植译, 神经网络原理，机械工业出版社，2006年
 - **Matin T. Hagan et al., Neural Network Design(神经网络设计), 机械工业出版社, 2005年**
 - Matin T. Hagan et al.著，戴葵等译, 神经网络设计，机械工业出版社，2002年.
 - **Fredric M. Ham and Ivica Kostanic, Principles of Neurocomputing for Science and Engineering (影印版), 机械工业出版社，2003年.**
 - Fredric M. Ham，Ivica Kostanic 著，叶世伟，王海娟译，神经计算原理，机械工业出版社，2007年.
 - **N. Cristianini and J. Shawe-Taylor. Introduction to Support Vector Machines. Cambridge, 2000.**
 - 李国正等译，支持向量机导论(An introduction to Support Vector Machines and other Kernel-based Learning Methods), 电子工业出版社，2004年3月
 - **I. Steinwart and A. Christmann. Support vector machines. Springer, New York, 2008.**

机器学习与统计理论书目 (2)

- 统计机器学习理论
 - Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar: **Foundations of Machine Learning**, MIT Press, 2012.
 - Kevin Patrick Murphy: **Machine Learning: a Probabilistic Perspective**, MIT Press, 2012.
 - L. Devroye, L. Györfi, and G. Lugosi. **A Probabilistic Theory of Pattern Recognition**. Springer, 1997.
 - Trevor Hastie et al., **The Elements of Statistical Learning(影印版)**, 2009.01
 - Trevor Hastie et al.著, 范明等译, 统计学习基础: 数据挖掘、推理与预测, 北京: 电子工业出版社, 2004年1月.
 - V. N. Vapnik. **Statistical Learning Theory**. Wiley, 1998.
 - V. N. Vapnik. **The Nature of Statistical Learning Theory**. Springer, 1995.
- 统计理论与应用统计分析
 - Larry Wasserman: **"All of Nonparametric Statistics"**, Springer 2007.
 - Larry Wasserman: **"All of Statistics: A Concise Course in Statistical Inference"**, Springer 2004.
 - R.A.Johnson and D.W.Wichern , **Applied Multivariate Statistics Analysis(5th Ed.)**, May, 2003.
 - 陆璇等译, 应用多元统计分析 (4th Ed), 北京: 清华大学出版社, 2001.
 - Daphne Koller and Nir Friedman: **"Probabilistic Graphical Models: Principles and Techniques"**, The MIT Press, 2009.
 - M. I. Jordan and C.M. Bishop, **An introduction to probabilistic graphical models**, University of California, Berkeley, 2003.

矩阵分析/计算与最优化 (3)

- 矩阵分析与计算
 - Roger A. Horn and Charles R. Johnson, **Matrix Analysis**
 - Gene H. Golub and Charles F. Van Loan: **“Matrix Computation” (4th Edition)**, Johns Hopkins University, 2012
 - 袁亚湘等译，矩阵计算，北京：科学出版社，2001.
 - 张贤达著，**矩阵分析与应用**，清华大学出版社，2004.9.
 - 方保镕、周继东、李医民等，**矩阵论**，清华大学出版社，2004.
- 优化理论
 - Stephen Boyd and Liewen Vandenberghe , **Convex Optimization**, Cambridge Univ. Press, 2004.
 - 陈宝林，**最优化理论与算法(第二版)**，清华大学出版社，2005年.
 - Dimitri P. Bertsekas: **“Convex Optimization Theory”**, Athena Scientific 2009.
 - David G. Luenberger: **“Optimization by Vector Space Methods”** , Wiley, 1969.
 - Dimitri P. Bertsekas: **“Constrained Optimization and Lagrangian Multiplier Methods”** , Academic Press, 1982.

机器学习中文教程

- 1. 于剑，机器学习——从公理到算法，清华大学出版社，2017
- 1. 周志华，机器学习，清华大学出版社，2016
- 2. 李航，统计学习方法，清华大学出版社，2012.

机器学习课程资源

- 1. Caltech
 - **Learning from data**
 - Yaser Abu-Mostafa
- 2. UBC
 - **Machine Learning**
 - Nando de Freitas
- 3. CMU
 - **Machine Learning**
 - Alex Smola
- 4. Stanford
 - **Machine Learning**
 - Andrew Y. Ng
- 5. MIT
 - **Statistical Learning Theory and Applications**
 - Tomaso Poggio
- 6. University of Wisconsin-Madison.
 - **Statistical Learning Theory**
 - Robert Nowak

Q / A

- Any Questions...