清华信息科学与技术国家实验室 计算机科学与人工智能研究部
TSINGHUA NATIONAL LABORATORY FOR INFORMATION SCIENCE AND TECHNOLOGY

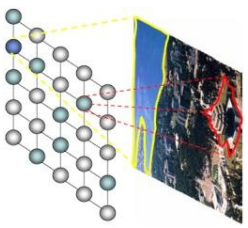
正则化贝叶斯学习
(Regularized Bayesian Learning)

课题简介

针对贝叶斯学习方法在考虑复杂问题的特定属性及领域知识方面存在的局限，本课题提出正则化贝叶斯学习的基本理论框架，研究其在结构化学习、隐含结构学习、最大间隔学习、高维空间稀疏学习、多模态信息融合、低秩矩阵分解、网络链接分析等方面的结合与应用；同时，研究准确高效的推理算法。本课题在学习理论与方法、问题提出与求解等方面都具有原创性。This project presents a new theoretical framework of regularized Bayesian learning and its extensions in structured learning, latent structure learning, max-margin learning, sparse learning in high dimensions, multi-modal data analysis, low-rank matrix factorization, link prediction, etc. We also develop fast and accurate learning algorithms.

代表性研究内容与成果描述

结构化学习（Structured Learning）



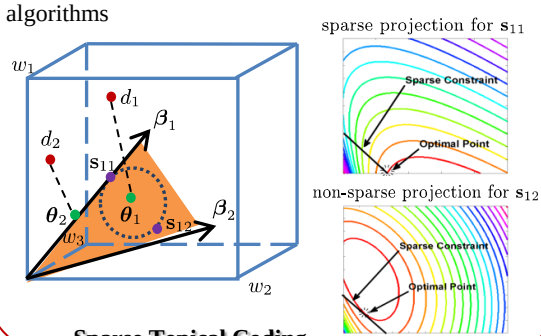
• 最大熵判别式马尔可夫网络为结构化学习提供了新理论和方法，并且有效考虑隐含结构以及高维空间稀疏学习等重要问题 Maximum entropy discrimination Markov networks provide a new theoretical framework for structured learning; We developed effective methods to consider latent structures and to learn sparse models in high dimensions;
• 提出对网页数据空间结构建模的二维条件随机场以及多层条件随机场模型，并且提出快速结构学习算法等 We developed 2D and hierarchical conditional random fields to model spatial layout structures of web data and fast algorithms to learn structures.

贝叶斯最大间隔学习（Bayesian Max-margin Learning）

• 提出正则化贝叶斯学习理论框架，将贝叶斯学习与最大间隔学习两个长期分离的领域自然有效地融合在一起，并提出若干有效的计算模型与算法 Regularized Bayesian learning provides a framework that naturally unifies Bayesian learning and max-margin learning, which have been largely isolated in machine learning and statistics.

参数化方法（Parametric Methods）

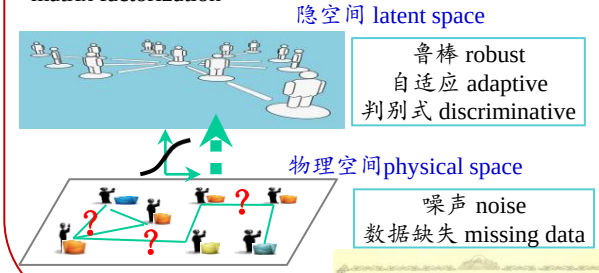
• 最大间隔话题模型及快速学习算法 MedLDA for learning topic structures and fast, scalable algorithms;
• 稀疏话题编码及在线学习 sparse topical coding for learning sparse topic structures and online learning algorithms



Sparse Topical Coding

非参数化方法（Nonparametric Methods）

•自适应无限支持向量机 Infinite SVM with infinite and adaptive clustering structures for complex data;
• 自适应无限隐含支持向量机及其在社会链接预测、矩阵低秩分解等问题的扩展及学习算法 Infinite latent SVM with adaptive latent features for link prediction, matrix factorization



Link Prediction

相关影响

• 机器学习顶级会议与杂志发表论文30余篇 More than 30 papers at major machine learning conferences and journals, including ICML, NIPS, UAI, JMLR, PAMI, etc.;
• 相关合作者受邀到CMU, Stanford, UC Berkeley, Princeton等著名学府做学术报告 Authors invited to give research talks at CMU, Stanford, UC Berkeley, Princeton, etc.;
• 相关博士论文“最大熵判别式马尔可夫网络：理论与应用”获中国计算机学会优秀博士论文奖 Thesis “Maximum Entropy Discrimination Markov Networks: Theory and Applications” won the China Computer Federation (CCF) distinguished Ph.D. thesis award in 2009.


获 奖 证 书
CERTIFICATE
朱军 博士
你的博士学位论文
“最大熵判别式马尔可夫网络：理论与应用”
被授予本学会2009优秀博士学位论文奖
特发此证
It is to certify that
Dr. Jun ZHU
whose doctoral dissertation
“Maximum Entropy Discrimination Markov Networks: Theory and Applications”
is awarded as the 2009 CCF outstanding dissertation.
China Computer Federation
January 30, 2010

Website: http://www.ml-thu.net/~jun/research.html

Supported by: National Key Project for Basic Research (973), National Natural Science Foundation of China, Tsinghua Initiative Scientific Research Program, Basic Research Foundation of Tsinghua National Lab for Information Science and Technology, and Tsinghua 221 Basic Research Plan for Young Faculties.