

统计学习理论

Statistical Learning Theory

MarkZZZ WeChat: MarkZZZ20XX

课程简介

统计学习理论 (Statistical Learning Theory) 是机器学习的数学基础, 研究学习算法在何种条件下、以何种速率能够从有限样本中泛化到未见数据。本课程以博士研究生为目标受众, 系统介绍 PAC 学习框架、VC 维理论、Rademacher 复杂度、均匀大数定律、结构风险最小化、支持向量机理论、Boosting、集成方法、在线学习、多臂赌博机、高维线性模型 (Lasso/稀疏恢复) 以及深度学习理论 (NTK、过参数化、双下降)。课程严格证明核心定理, 培养学生阅读前沿论文的数学素养与独立研究能力。

适合人群

- 统计学、计算机科学、数学方向的博士研究生
- 希望深入理解机器学习理论基础的研究人员
- 对泛化理论、高维统计或在线优化有研究兴趣者

前置知识

- **概率论:** 测度论基础、集中不等式 (Hoeffding, Azuma-McDiarmid)、大数定律
- **线性代数:** 矩阵分析、谱理论、正定矩阵
- **凸优化:** 对偶理论 (KKT 条件)、梯度方法
- **机器学习基础:** 监督/无监督学习基本概念、经验风险最小化

1 课程内容总览

| 讲次 | 主题 | 内容概要 |
|----|----------------|--|
| 1 | PAC 学习框架 | 有限假设类、可实现性与不可知性、一致性收敛、样本复杂度 $O(\ln \mathcal{H} /\epsilon^2)$ |
| 2 | VC 维理论 | Shattering、Vapnik-Chervonenkis 维、增长函数、Sauer-Shelah 引理完整证明、VC 推广界 |
| 3 | Rademacher 复杂度 | 定义与基本性质、Ledoux-Talagrand 收缩引理、与 VC 维的关系、数据依赖泛化界 |
| 4 | 均匀大数定律 | 对称化引理、Ghost 样本方法、Pollard 覆盖数与熵积分、Dudley 定理 |

| 讲次 | 主题 | 内容概要 |
|----|-------------|--|
| 5 | 结构风险最小化 | 偏差-方差权衡、MDL 原理、最小描述长度与贝叶斯先验的等价性、PAC-Bayes 界 |
| 6 | 支持向量机 | 间隔理论、RKHS 与 Mercer 定理、SVM 对偶完整推导、基于间隔的泛化界 |
| 7 | Boosting 理论 | AdaBoost 指数损失最小化、边界最大化定理、弱学习到强学习的等价性 |
| 8 | 集成方法理论 | Bagging 方差减少的严格证明、随机森林一致性 (Breiman-Biau)、偏差-方差分析 |
| 9 | 在线学习 | Regret 界、加权多数算法、Hedge 算法、跟随正则化领导者 (FTRL) 理论 |
| 10 | 多臂赌博机 | UCB 算法与遗憾界、Lai-Robbins 下界、Thompson 采样的贝叶斯遗憾分析 |
| 11 | 高维线性模型 | Lasso 相合性与预测误差、RIP 条件、SLOPE 自适应估计、稀疏恢复信息论下界 |
| 12 | 深度学习理论 | Neural Tangent Kernel、过参数化隐式正则、记忆容量界、双下降现象理论 |

2 主要参考文献

2.1 经典教材

- **Shalev-Shwartz & Ben-David**, *Understanding Machine Learning: From Theory to Algorithms*, Cambridge University Press, 2014。
最适合作为本课程主教材，PAC 学习、VC 维、Rademacher 复杂度均有系统介绍。
- **Vapnik, V. N.**, *Statistical Learning Theory*, Wiley, 1998。
统计学习理论的奠基著作，VC 理论原始来源。
- **Kearns & Vazirani**, *An Introduction to Computational Learning Theory*, MIT Press, 1994。
PAC 学习的经典导论，计算视角鲜明。
- **Bartlett & Mendelson (2002)**, “Rademacher and Gaussian Complexities: Risk Bounds and Structural Results”, *JMLR* 3, 463–482。
Rademacher 复杂度理论的核心参考文献。
- **Boucheron, Lugosi & Massart**, *Concentration Inequalities: A Nonasymptotic Theory of Independence*, Oxford University Press, 2013。
集中不等式的权威参考，涵盖 Efron-Stein、对数 Sobolev 不等式等。
- **Wainwright, M. J.**, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, Cambridge University Press, 2019。
高维统计 (Lasso、稀疏恢复、RIP) 的最佳现代参考。
- **Cesa-Bianchi & Lugosi**, *Prediction, Learning, and Games*, Cambridge University Press, 2006。