

李健

手机: (+86) 130-2005-8233 邮箱: lijian9026@ie.ac.cn

学术主页: lijian.ac.cn Github: github.com/superlj666

学历: 博士研究生 预计毕业时间: 2021 年 1 月 研究方向: 大规模半监督机器学习

教育背景

2015.09 - 2021.01	中国科学院大学-信息工程研究所-工学博士	机器学习
2011.09 - 2015.07	东北大学-软件学院-工学学士	软件工程 (英语)

研究方向及兴趣

大规模半监督机器学习：理论分析与算法设计

针对传统机器学习方法应用于大规模、半监督数据存在的理论及算法瓶颈, 研究兼具理论保证与可伸缩性的学习方法。具体来说, 结合随机优化、随机映射、分布式等大规模优化方法, 使用复杂度分析工具进行泛化误差分析, 并引入无标签数据降低部分误差, 最终指导高效算法的设计。已有工作概况:

- 理论: 使用拉德马赫复杂度、积分算子等分析工具, 对学习算法进行泛化误差分析, 从而指导算法设计。
- 准则: 基于理论分析制定简洁易实现的模型选择准则, 提高超参选择、模型设计的效率。
- 算法: 使用随机优化、随机映射、分布式等手段, 设计适用于大规模数据的高效可伸缩算法。

已发表论文

- [Multi-Class Learning using Unlabeled Samples: Theory and Algorithm.](#)
Jian Li, Yong Liu, Rong Yin, Weiping Wang.
In Proceedings of the 28th International Joint Conference on Artificial Intelligence (**IJCAI 2019**). To appear.
- [Approximate Manifold Regularization: Scalable Algorithm and Generalization Analysis.](#)
Jian Li, Yong Liu, Rong Yin, Weiping Wang.
In Proceedings of the 28th International Joint Conference on Artificial Intelligence (**IJCAI 2019**). To appear.
- [Multi-Class Learning: From Theory to Algorithm.](#)
Jian Li, Yong Liu, Rong Yin, Hua Zhang, Lizhong Ding, Weiping Wang.
Advances in Neural Information Processing Systems 31 (**NeurIPS 2018**).
- [Efficient kernel selection via spectral analysis.](#)
Jian Li, Yong Liu, Hailun Lin, Yinliang Yue, Weiping Wang.
In Proceedings of the 26th International Joint Conference on Artificial Intelligence (**IJCAI 2017**).

预印论文

- [Distributed Learning with Random Features.](#)
Jian Li, Yong Liu, Weiping Wang.
arXiv preprint arXiv:1906.03155, 2019. (Submitted to NeurIPS 2019)
- [Efficient Cross-Validation for Semi-Supervised Learning.](#)
Yong Liu, **Jian Li**, Guangjun Wu, Lizhong Ding, Weiping Wang.
arXiv preprint arXiv:1902.04768, 2019.
- [Max-Diversity Distributed Learning: Theory and Algorithms.](#)
Yong Liu, **Jian Li**, Weiping Wang.
arXiv preprint arXiv:1812.07738, 2018.

专业技能

- **机器学习理论:** 深入学习理解核方法 (SVM、KRR 等) 相关泛化分析理论, 熟悉半监督学习近似泛化理论、大规模核方法近似泛化理论.
- **机器学习算法:** 熟悉常见机器学习优化算法 (GD、SGD、牛顿法等)、了解算法收敛性分析, 有丰富的优化算法设计实现经验.
- 熟悉经典机器学习开发应用流程及常用工具包 (scikit-learn, Pandas, NumPy, matplotlib 等), 了解 CV、NLP 领域经典深度学习模型及常用开发框架.
- 熟悉 Linux 平台, 熟练使用 Python、Matlab、Java 等开发语言及开发工具.
- 良好的数理统计基础, 逻辑分析能力.
- 良好的英文文献阅读能力 (CET-6 517), 流利的英语口语交流能力.

项目经历

- 2017.04-2017.12 **大规模交互式情报分析研判平台** 中科院信工所, 后端算法开发工程师
- 负责交互式大规模机器学习算子的设计及开发工作.
 - 负责交互式大规模图计算 (群体发现及分析、重点人发现及分析) 模块的设计及开发工作.
 - 负责人物画像模块的开发工作 (通过构建简单知识图谱、规则匹配实现).
- 2016.07-2017.03 **YJ 子系统** 中科院信工所, 后端算法开发工程师
- 负责大规模数据开发集群的搭建及维护.
 - 负责大规模机器学习算子、大规模图计算 (社区发现、传播溯源、行为分析) 等模块的设计及开发.
 - 负责对已有算法进行优化以适用于大规模环境并达到性能指标.
- 2015.01-2015.09 **剑与魔法手游** 北京游奕互动有限责任公司, U3D 交互逻辑开发工程师
- 负责角色状态转换的设计与开发工作.
 - 负责部分 UI 交互逻辑的开发工作.

获奖情况

- 2018 年博士研究生国家奖学金, 国家级. 30000 元.
- 2019 年中国科学院院长优秀奖, 校级. 5000 元.
- 2018 年信息工程研究所所长优秀奖, 所级. 3000 元.
- 2018-2019 学年中国科学院大学三好学生, 校级.
- 2017-2018 学年中国科学院大学三好学生, 校级.
- 2018 年中科院信工所第二研究室优秀学生, 所级. 2000 元.
- 2017 年中科院信工所第二研究室优秀学生, 所级. 2000 元.