

# 李健

☎: (+86) 130-2005-8233    ✉: me@lijian.ac.cn

🏠: lijian.ac.cn    📄: Google Scholar    🌐: github.com/superlj666

学历: 博士研究生    预计毕业时间: 2020 年 6 月    研究方向: 大规模半监督机器学习

## 教育背景

2015.09 - 2020.06	中国科学院-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 Rademacher 复杂度、积分算子等分析工具, 对算法进行泛化性能分析, 进而指导算法设计.
- 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.

### 神经网络可解释性

基于随机傅里叶特征等非线性映射及 Rademacher 复杂度, 从核方法的角度理解神经网络.

## 已发表论文 (🔗 GOOGLE SCHOLAR PROFILE)

- [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**).
- [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**).
- [Multi-Class Learning: From Theory to Algorithm.](#)  
**Jian Li**, Yong Liu, Rong Yin, Hua Zhang, Lihong 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

- [Automated Spectral Kernel Learning.](#) (Submission in AAAI)  
**Jian Li**, Yong Liu, Weiping Wang.  
arXiv preprint arXiv:1909.04894, 2019.
- [Learning Vector-valued Functions with Local Rademacher Complexity.](#) (Submission in TPAMI)  
**Jian Li**, Yong Liu, Weiping Wang.  
arXiv preprint arXiv:1909.04883, 2019.
- [Distributed Learning with Random Features.](#)  
**Jian Li**, Yong Liu, Weiping Wang.  
arXiv preprint arXiv:1906.03155, 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](#). (Submission in AAAI)  
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、神经网络可解释性.
- **机器学习模型**: 深入理解核方法、熟悉常见机器学习模型、了解经典深度神经网络模型.
- **机器学习算法**: 熟悉常见机器学习优化算法 (一阶梯度方法、牛顿法等)、了解相关算法收敛性分析.
- **机器学习开发环境**:
  - 熟练使用 Linux 平台开发环境及常用命令.
  - 熟练使用机器学习开发流程及常用工具包 (scikit-learn, Pandas, NumPy, matplotlib 等).
  - 熟悉经典深度学习开发框架 (Pytorch、TensorFlow).
  - 熟悉大规模数据挖掘算法开发框架 (Spark), 有大规模算法开发经验.
- **编程语言**: 掌握 Python、Matlab, 熟悉 C/C++、Java 等编程语言.
- **数理基础**: 良好的数理统计基础, 逻辑分析能力.
- **外语水平**: 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, 数据挖掘算法开发
- 负责 Spark 平台上常见机器学习算法 (逻辑斯蒂回归、线性回归、k 近邻等模型) 的设计及实现.
  - 负责将底层算法应用到项目业务需求中, 对大规模数据进行分析.
  - 负责对平台算法进行优化以适用于大规模环境并达到性能指标.
- 2016.07-2017.03      **邮件分析子系统**      中科院信工所, 大规模图计算开发
- 负责大规模数据开发环境 (Spark 集群) 的搭建及维护.
  - 负责大规模图计算方法 (社区发现、关键节点发现、传播溯源等) 的设计及实现.
  - 负责将底层图计算对大规模图进行分析 (群体发现、传播溯源、行为分析、重点人发现等) .
- 2014.12-2015.09      **剑与魔法手游**      北京游奕互动, U3D 交互逻辑开发
- 负责角色状态转换机的设计与开发工作.
  - 负责部分 UI 交互逻辑的开发工作.
  - 优化游戏打包压缩效率、服务器与客户端通信效率.

## 获奖情况

---

- 拟获国科大博士研究生国际合作培养计划 (2%, USD \$22,800), 用于赴斯坦福大学联合培养 12 个月. 中国科学院大学. 2019.
- 中国科学院院长优秀奖 (Top 1%, RMB ¥5,000), 中国科学院. 2019.
- 博士研究生国家奖学金 (Top 2%, RMB ¥30,000), 教育部. 2018.
- 信息工程研究所所长优秀奖 (Top 10%, RMB ¥2,000), 中科院信工所. 2018.
- 国科大”三好学生”荣誉称号, 中国科学院大学. 2018.
- 国科大”三好学生”荣誉称号, 中国科学院大学. 2019.
- 第二研究室优秀学生, 中科院信工所. 2017.
- 第二研究室优秀学生, 中科院信工所. 2018.