

许坚

(+86)18851070119

msgsxj@gmail.com

教育经历

南京大学数学系 概率论与数理统计学硕士 2017/09 - 2020/06

主修课程：高等概率论、高等数理统计学、随机过程、数值代数、多元迭代分析、一阶优化方法

扬州大学数科院 统计学学士 2013/09 - 2017/06

主修课程：概率论、数理统计、时间序列分析、多元统计分析、回归分析、应用随机过程

技能

| | |
|-------|-----------------------------|
| 写作和演示 | LaTeX |
| 代码 | python、numpy、pandas、pytorch |
| 医疗 | 熟悉各类医学影像软件、vtk |
| 数学优化 | 理解各类最优化算法，能够结合具体问题快速实现 |

研究方向

机器学习中的**最优化**，包括特征选择、稀疏学习。通过对最优化领域经典算法（数值优化等）的不断学习、对机器学习模型及对应优化问题的不断了解，期望碰到问题能设计出高效的优化算法。

机器学习中的**概率论**，包括极限理论部分，即以尾概率事件的不等式去度量机器学习算法学习失败的概率。

工作经历

苏州智慧医疗创新中心 助理研究员 2020/08-至今

基于联邦学习的 AI 模型融合平台。该平台能够对相同模型结构的基于 pytorch 的 AI 模型进行模型融合，以用于评估心肌纤维化程度的 AI 模型为例，该平台除了联邦学习保护病人数据隐私的特性外，还会根据基于 Rademacher 复杂度的上界去评估每个参与融合的子模型的泛化误差，作为模型融合的初始起点。（该项目已申请两个专利，一作）

基于遗传算法的符号回归。这是基于 changwei loh(罗章维、创新中心首席科学家) 的一项工作<https://arxiv.org/abs/1908.10673> 的探索，该领域的目标为从数据中自动找公式。与之前的结果相比，新的方法能够找到隐函数，并能够处理该领域的金标准——双摆问题。（准备投递期刊）

福特中国 数据科学家 2019/08 - 2019/10

上海福特拼车项目，目标是对拼车价格给出一个定价参考。该项目主要有两个难点，一是订单的数据量很小，无法给出可靠性较高的定价参考，在针对性地做了一些实验设计的工作后，实验车队的日订单量破千的同时能够去简单区分价格敏感型用户；第二个难点是如何选择适合的模型去建模，这里我们为了避免引入过多的先验假设，选用广义线性模型为基本元素来搭建模型，目标为日营收额最大化，并选用坐标下降法解对应的大规模约束优化问题。

竞赛经历

| | | |
|--|-------------|---------|
| 阿里巴巴人工智能对抗算法赛 | 排名:124/2519 | 2019/02 |
| 熟悉了 cleverhans.attacks 这一攻击模，读了一些现有的方法并使用了它们。 | | |
| 阿里巴巴智能云上编程大赛 | 排名:14/1022 | 2018/07 |
| 使用 PAI 平台进行建模，简单考虑了一些时间、空间上的统计特征。 | | |
| 全国研究生数学建模 | 三等奖 | 2017/09 |
| 负责从文中题目到模型建立的部分，以及对应优化问题的处理，最终优化算法选用 Nelder-Mead 算法。 | | |
| 全国大学生数学竞赛 (数学类) | 三等奖 | 2016/10 |