

LI, Geng | 李耕

哈尔滨工业大学，计算学部

Mobile: 86-13030040197 | Email: ligeng@stu.hit.edu.cn

教育经历

哈尔滨工业大学:

学术硕士研究生: 计算科学与技术专业, 计算学部.

2021 - 2023

• 核心课程:

1. 机器学习&算法: 机器学习理论和算法 (91); 高级算法设计与分析 (90)
2. 学术写作: 学术写作及学术规范 (93.2)

本科: 计算科学与技术专业, 计算学部 & 英才学院.

2017 - 2021

• 排名: 3/50

• 核心课程:

1. 数学: 微积分 (97); 线性代数与几何分析 (95).
2. 计算理论: 数理逻辑 (98.2); 概率论与数理统计 (92); 形式语言与自动机 (96)
3. 算法 & 软件: 算法设计与分析 (91); 软件构造 (93.7); 数据库系统 (93); 计算机网络 (95); 数据挖掘 (97)
4. 语言: 英语国际会议 (98.8); 国际交流英语 (96).
5. 约翰霍普金斯交流访问课程: Machine Learning(A); Randomized & Big Data Algorithms(A-)

荣誉, 获奖 & 奖学金

- 国家励志奖学金 (2018, 2019, 2020)
- 一等人民奖学金 (2018, 2019, 2020)
- 英才学院十佳优秀毕业生 (2021)
- 校优秀毕业生 (2021)
- 校级优秀学生 (2017-2018, 2019-2020)
- 校级优秀团干部 (2022)
- 2021阿里云基础设施供应链大赛排名23/1731 (1. 3%)
- 2020全国大学冰壶人工智能挑战赛二等奖

发表工作

• “EEML: Ensemble Embedded Meta-learning”

(WISE 2022接收, arxiv 预印版: <https://arxiv.org/abs/2206.09195>)

作者: Geng Li, Boyuan Ren, Hongzhi Wang

• “FL-AGCNS: Federated Learning Framework for Automatic Graph Convolutional Network Search”

(TKDE审核中, 第一轮审稿意见与反馈已完成, arxiv 预印版: <https://arxiv.org/abs/2104.04141>)

作者: Chunnan Wang, Bozhou Chen, Geng Li, Hongzhi Wang

研究经历

项目 I: ASML: 基于模块状态的架构相关元学习

Start: Jan 2022

导师: 王宏志, 海量数据计算研究中心主任, 哈尔滨工业大学

- 提出一种开销可控的元学习算法适用于计算资源受限的设备 (GPU显存较小)
- 设计状态管理器用于针对不同任务动态调整模型组件学习策略。
- 执行相关对比实验证明算法有效性。

项目 II: EEML: 基于集成嵌入元学习

Sep 2020 – Jun 2021

导师: 王宏志, 海量数据计算研究中心主任, 哈尔滨工业大学

- 提出一种针对小样本任务的元学习算法;
- 利用梯度信息表征每个仅包含少量样本的任务数据集;
- 设计任务嵌入聚类、元训练、元测试过程来考虑异质任务间的联系与差异;
- 执行相关实验验证算法有助于增强现有基于梯度元学习 (MAML) 在回归和图像分类任务上的表现性能。
- 工作已被 2022 Web Information Systems Engineering接收。

导师: 王宏志, 海量数据计算研究中心主任, 哈尔滨工业大学

- 学习联邦学习、遗传算法、强化学习和神经网络架构搜索 (NAS)
- 实现了论文中图神经网络搜索架构及遗传算法迭代更新
- 实现强化学习算法和随机采样算法作为对比实验
- 实现内部消融实验用于分析和支持算法优越性
- 工作在IEEE Transactions on Knowledge and Data Engineering (TKDE)审稿中, 已经历第一轮审稿意见及反馈。

证书

- **CET6:** 总分 505 | 听力 162 | 阅读 191 | 写作 & 翻译 152
- **CET4:** 总分 570 | 听力 218 | 阅读 196 | 写作 & 翻译 156

项目经历

- **AI4DB: GaussDB数据库旋钮效果的快速评估机制.**

Start: May 2022

导师: 王宏志, 海量数据计算研究中心主任, 哈尔滨工业大学

利用元学习针对小样本场景的适用性, 在数据库负载动态变化的情况下增强数据库旋钮评估模型的泛化能力, 以降低负载迁移后模型再训练的成本, 并提高最终旋钮评估性能。

- **Kaggle: BirdCLEF 2022鸟类叫声识别比赛.**

Feb – May 2022

担任队长.

设计了两个阶段模式 (检测和分类) 以从声音片段中识别鸟类物种;

应用 Audio Spectrogram Transformer (AST)、Vision Transformer (ViT)、AlexNet、CNN-RNN 等网络作为骨干网络, 检测片段中是否存在鸟声以及进行分类。

- **天池竞赛: 2021阿里云基础设施供应链大赛**

Oct 2021 – Jan 2022

担任队长.

提出两部分 (预测&决策) 解决方案, 帮助阿里云 (云主机提供商) 维持最佳利润供应链;

预测: 利用 LSTM、ARIMA、DeepAR 时序模型预测未来需求;

决策: 设计动态安全库存策略以支持决策。

- **课程项目: 机器学习相关课程实验 (约翰霍普金斯交换项目期间) .**

系统学习并复现经典机器学习算法: Logistic 回归、决策树 (ID3)、KNN、GMM、K-means、MLP、SVM 并在公开benchmark上进行测试。

技能

- **编程语言:** Python, Java, C.
- **机器学习 & 深度学习:** Pytorch, Pandas, Numpy, Learn2Learn, Scikit-learn, .
- **其他软件工具:** Latex, Linux, Git.