

# 沈彧

🔗 salty-fish-97.github.io    🌐 github.com/salty-fish-97

☎ (+86) 18930101735    ✉ shenyu@pku.edu.cn

📍 北京市海淀区颐和园路 5 号 1065 室

📅 1997 年 5 月 16 日出生于上海市虹口区

## 机器学习系统工程师

**简介** 我是北京大学计算机学院网络所四年级博士生，师从崔斌教授，主要研究方向是**黑盒优化**，**自动化机器学习系统**以及**算法**。

**研究兴趣** 博士期间，我的研究工作涵盖了很多方面的主题，主要包括：**黑盒优化**，**自动化机器学习**，**集成学习**，**迁移学习**，**自动化机器学习系统**，**图神经网络**等。目前，我对自动化机器学习非常感兴趣，包括各类子领域（自动化超参数优化，自动化特征工程，网络结构搜索等），及其在业务场景上的应用（数据库，推荐，软件性能优化等）。

## 🎓 教育背景

2019 年 9 月 – 2024 年 6 月(expected)	博士在读，计算机软件与理论，计算机学院，北京大学，北京 导师：崔斌教授
20015 年 9 月 – 2019 年 6 月	本科，计算机软件方向，信息科学技术学院，北京大学，北京 毕业设计题目：“使用声明式语言的自动化深度学习系统” 导师：崔斌教授，学院优秀毕设

## 📖 学术论文

- **Shen, Y.**, Zheng, J., Li, Y., Zhang, W., Yao, P., Li, J., Yang, S., Liu, J., and Cui, B. (2023). ProxyBO : Accelerating Neural Architecture Search via Bayesian Optimization with Zero-cost Proxies. To appear in AAAI. **CCF-A**
- **Shen, Y.**, Lu, Y., Li, Y., Tu, Y., Zhang, W., and Cui, B. (2022). DivBO : Diversity-aware CASH for Ensemble Learning. NeurIPS. **CCF-A**
- Li, Y., **Shen, Y.**, Zhang, W., Zhang, C., and Cui, B. (2022). VolcanoML : speeding up end-to-end AutoML via scalable search space decomposition. VLDBJ. **CCF-A**
- Zhang, W., Lin, Z., **Shen, Y.**, Li, Y., Yang, Z., and Cui, B. (2022). Deep and Flexible Graph Neural Architecture Search. ICML. **CCF-A**
- Zhang, W., Sheng, Z., Yang, M., Li, Y., **Shen, Y.**, Yang, Z., and Cui, B. (2022). NAFS : A Simple yet Tough-to-beat Baseline for Graph Representation Learning. ICML. **CCF-A**
- Li, Y., **Shen, Y.**, Jiang, H., Zhang, W., Yang, Z., Zhang, C., and Cui, B. (2022). TransBO : Hyperparameter Optimization via Two-Phase Transfer Learning. KDD. **CCF-A**
- Li, Y., **Shen, Y.**, Jiang, H., Bai, T., Zhang, W., Zhang, C., and Cui, B. (2022). Transfer Learning based Search Space Design for Hyperparameter Tuning. KDD. **CCF-A**
- Zhang, W., **Shen, Y.**, Lin, Z., Li, Y., Li, X., Ouyang, W., Tao, Y., Yang, Z., and Cui, B. (2022) PaSca : a Graph Neural Architecture Search System under the Scalable Paradigm. **WWW Best Student Paper**. **CCF-A**
- Li, Y., **Shen, Y.**, Jiang, H., Zhang, W., Li, J., Liu, J., Zhang, C., and Cui, B. (2022). Hyper-Tune : Towards Efficient Hyperparameter Tuning at Scale. VLDB. **CCF-A**
- Zhang, W., Jiang, Y., Li, Y., Sheng, Z., **Shen, Y.**, Miao, X., Wang, L., Yang, Z., and Cui, B. (2021). ROD : Reception-aware Online Distillation for Sparse Graphs. KDD. **CCF-A**
- Li, Y., **Shen, Y.**, Zhang, W., Chen, Y., Jiang, H., Liu, M., Jiang, J., Gao, J., Wu, W., Yang, Z., Zhang, C., and Cui, B. (2021). OpenBox : A Generalized Black-box Optimization Service. KDD. **CCF-A**

- > Zhang, W., Yang, Z., Wang, Y., **Shen, Y.**, Li, Y., Wang, L., and Cui, B. (2021). Grain : Improving Data Efficiency of Graph Neural Networks via Diversified Influence Maximization. VLDB. **CCF-A**
- > Li, Y., **Shen, Y.**, Zhang, W., Jiang, J., Ding, B., Li, Y., Zhou, J., Yang, Z., Wu, W., Zhang, C., and Cui, B. (2021). VolcanoML : Speeding up End-to-End AutoML via Scalable Search Space Decomposition. VLDB. **CCF-A**
- > Zhang, W., **Shen, Y.**, Li, Y., Chen, L., Yang, Z., and Cui, B. (2021). ALG : Fast and Accurate Active Learning Framework for Graph Convolutional Networks. SIGMOD. **CCF-A**
- > Li, Y., **Shen, Y.**, Jiang, J., Gao, J., Zhang, C., and Cui, B. (2021). MFES-HB : Efficient Hyperband with Multi-Fidelity Quality Measurements. AAAI. **CCF-A**

## ☰ 相关技能

编程技能： **Python**,  $\text{\LaTeX}$ , C++, Java.

机器学习： 掌握 贝叶斯优化算法，遗传算法，启发式算法

掌握 传统机器学习算法，深度学习算法

熟悉计算机视觉算法，自然语言处理算法

了解最优化理论（凸优化）

## </> 实习/项目经历

至今	项目负责人, 北京大学 & 字节, 校企合作项目
2022 年 4 月	<ul style="list-style-type: none"> <li>&gt; 2022 年 4 月 - 2022 年 7 月, 调研较为通用的迁移学习场景框架, 并总结合理可行的执行方案。</li> <li>&gt; 2022 年 8 月 - 至今, 研究针对 <i>Spark</i> 优化场景的通用迁移学习参数优化算法; 主要思想是通过将专家知识以及过去的验证历史融入当前的参数优化流程中, 在给定较少的调参预算下, 大幅度降低 <i>Spark</i> 任务的执行成本。</li> <li>&gt; 算法在业务场景初步测试获得降低默认参数近 60% 的成本。</li> </ul> <div>黑盒优化, 软件参数优化 <span>Spark</span> <span>Python</span></div>
至今	项目负责人, 北京大学 & 中兴, 校企合作项目
2022 年 4 月	<ul style="list-style-type: none"> <li>&gt; 2022 年 4 月 - 2022 年 7 月, 调研针对深度学习的自动化超参数优化算法; 通过权衡在实际场景中的收益 - 成本的关系, 以联合算法选择以及超参数优化 (CASH) 作为后续主要研究以及交付方向。</li> <li>&gt; 2022 年 8 月 - 至今, 研究针对深度学习的 <i>CASH</i> 方法; 主要思想是通过引入贝叶斯优化以及空间切分策略, 取代人力调参, 并搭建通用的底层框架, 支持各类深度学习任务的调参。同时, 基于各类任务的特性, 添加早停机制、模型集成等策略。</li> </ul> <div>AutoML <span>自动化超参数优化</span> <span>神经网络结构搜索</span> <span>Python</span></div>
2021 年 10 月	项目负责人, 北京大学 & 快手, 校企合作项目
2020 年 10 月	<ul style="list-style-type: none"> <li>&gt; 2020 年 10 月 - 2021 年 5 月, 研究大规模自动化超参数优化算法; 主要思想是通过结合多精度验证, 异步并行的思想实现大规模自动化超参数优化, 在不影响算法收敛结果的前提下, 提升验证机器的工作效率。</li> <li>&gt; 2021 年 5 月 - 9 月, 研究使用零消耗代理的神经网络结构搜索算法; 主要思想是结合零消耗代理与贝叶斯优化算法的优点, 缓解贝叶斯优化算法前期因缺乏有效验证无法给出优秀配置的问题, 并大幅提高搜索效率。</li> <li>&gt; 大规模自动化超参数优化算法在业务场景提升下游推荐任务 AUC 一个百分点。</li> </ul> <div>AutoML <span>自动化超参数优化</span> <span>神经网络结构搜索</span> <span>Python</span></div>

2021 年 6 月	MindWare 端到端自动化机器学习系统开发, 北京大学, 实验室项目
2020 年 10 月	<ul style="list-style-type: none"> <li>独立完成系统各技术栈代码, 包含底层各机器算法实现, 中间层自动化搜索算法实现, 以及顶层系统模块封装。</li> </ul> <div>AutoML Python</div>
至今	OpenBox 黑盒优化服务开发, 北京大学, 实验室项目
2020 年 10 月	<ul style="list-style-type: none"> <li>负责单目标无约束问题的黑盒优化算法实现。</li> <li>负责服务可视化接口设计。</li> <li>负责后续维护以及新功能扩展。</li> </ul> <div>AutoML 黑盒优化 Python</div>
2019 年 6 月	使用声明式语言的自动深度学习系统, 北京大学, 本科毕业设计
2019 年 3 月	<ul style="list-style-type: none"> <li>设计声明式语言, 提供简便的任务定义接口;</li> <li>搭建自动化深度学习框架, 实现大量深度学习模型, 并通过 AutoML 算法实现自动搜索;</li> <li>在图像分类问题上效果显著优于现有系统 Auto-Keras;</li> <li>获得学院优秀毕设。</li> </ul> <div>AutoML 深度学习 Python</div>

## 语言

英语: 阅读 ● ● ● ● ● 4 级: 625  
 听力 ● ● ● ● ● 6 级: 533  
 口语 ● ● ● ● ○

## 奖励与荣誉

2022 年 11 月 北京大学国家奖学金 (<1.6%)  
 2022 年 11 月 北京大学三好学生标兵 (<1.6%)  
 2021 年 12 月 第四届 CCF 软件开源大赛自由创新组 OpenBox 获赛道**唯一特等奖**  
 2021 年 11 月 北京大学兴业银行奖学金 (<7.5%)  
 2021 年 11 月 北京大学三好学生 (<12%)  
 2021 年 10 月 2021 QQ 浏览器 AI 算法大赛暨 2021 CIKM AnalytiCup 自动化超参数优化赛道初赛与决赛 **冠军**

## 兴趣爱好

体育: 游泳, 乒乓  
 艺术: 小提琴