

姚远

手机：13751062691 · 邮箱：yaoyuan.hitsz@gmail.com

[Homepage](#) · [Google scholar](#) · [Github](#)



基本信息

- 所在城市：北京市；出生日期：1994 年 5 月 24 日

教育背景

哈尔滨工业大学，博士（直博）2015.09 - 2021.06

- 导师：叶允明教授，李旭涛教授；专业：计算机软件与理论；学分绩点：3.27/4.0

西北民族大学，学士2011.09 - 2015.07

- 专业：计算机科学与技术；专业排名：1/79

工作与访问经历

百度，高级算法工程师2021.08 - 至今

- 部门：搜索策略部

南洋理工大学，国家公派联合培养博士研究生2019.10 - 2020.07

- 导师：Sinno Jialin Pan 副教授

香港科技大学，访问实习生2017.10 - 2018.04

- 导师：杨强教授，张宇副教授（目前任职于南方科技大学）

研究方向

- 异构迁移学习/多模态学习，深度学习，自然语言处理，信息检索

项目经历

优化百度官网搜索排序机制，提升百度官网搜索结果的排序合理率，现将主要项目经历总结如下：

1. 官网寻址 RRT (Recall Ranking Trigger) 框架

- 项目内容：该项目旨在提升官网首页排序合理率，主要包括离线召回、离线排序和在线触发三个阶段。离线召回阶段负责从百亿级数据中召回候选官网 query-url，离线排序阶段负责筛选高置信官网 query-url，在线触发阶段负责结合用户 query 对合适的官网 url 进行优先展现。
- 主要负责：负责离线召回和离线排序阶段的搭建。离线召回阶段通过设计特定的规则从点击日志数据和互联网备案数据中产出大量候选官网 query-url；离线排序阶段通过设计并抽取 query-url 官需满足类特征，构建 GBDT 二分类模型，筛选高置信官网 query-url。
- 项目成果：该项目已在百度搜索引擎中应用，并在官网首页排序合理率指标上持续超越谷歌。

2. query 官网需求强度预测模型

- 项目内容：该项目旨在建设官网内页排序机制的触发信号，用于提升官网内页排序合理率。
- 主要负责：负责利用 query 官需类样本在 12 层 ERNIE 通用预训练模型上进行参数微调，实现通用预训练模型到 query 官网需求强度预测模型的知识迁移。
- 项目成果：该项目已在百度搜索引擎中应用，同时也作为一维特征在官网寻址 RRT 框架中应用。

研究经历

研究多源异构数据的知识迁移问题，提升多源异构数据的可迁移性，现将代表性研究成果总结如下：

1. 基于柔性迁移网络 (Soft Transfer Network, STN) 的异构迁移学习方法研究

- 应用场景：跨语言文本分类、跨模态图像识别。
- 模型创新：首次在异构迁移学习领域利用目标领域无类标样本的软标签信息约减源领域和目标领域条件分布的差异，并利用非线性变换学习最优的公共子空间。

2. 基于条件加权对抗网络（Conditional Weighting Adversarial Network, CWAN）的多源异构迁移学习方法研究
- 应用场景：跨语言文本分类、跨模态图像识别。
 - 模型创新：首次利用深度对抗学习技术处理多源异构迁移学习问题。此外，该方法引入条件加权机制，不仅衡量了每个源领域的重要程度，而且约减了所有源领域和目标领域条件分布的差异。
3. 基于判别式分布对齐（Discriminative Distribution Alignment, DDA）框架的异构迁移学习方法研究
- 应用场景：跨语言文本分类、跨模态图像识别。
 - 模型创新：其首次将分类器自适应约束，分布对齐约束和判别式嵌入约束整合至统一的框架中，并基于此框架嵌入两个不同的损失函数，构建出两个不同的异构迁移学习方法。
4. 基于异构迁移学习的低分辨率图像分类问题研究
- 应用场景：低分辨率图像分类。
 - 模型创新：首次将低分辨率图像分类问题建模成异构迁移学习问题，并提出判别式联合分布自适应方法，其不仅加大了不同类图像之间的距离，而且约减了高低分辨率图像分布之间的差异。

学术论文

- Yuan Yao, Yu Zhang, Xutao Li, and Yunming Ye. Heterogeneous Domain Adaptation via Soft Transfer Network. **ACM MM’19 (CCF A)**.
- Yuan Yao, Xutao Li, Yu Zhang, and Yunming Ye. Multi-source Heterogeneous Domain Adaptation with Conditional Weighting Adversarial Network. **IEEE TNNLS’21 (CCF B)**.
- Yuan Yao, Yu Zhang, Xutao Li, and Yunming Ye. Discriminative Distribution Alignment: A Unified Framework for Heterogeneous Domain Adaptation. **PR’20 (CCF B)**.
- Yuan Yao, Xutao Li, Yunming Ye, Feng Liu, Michael K. Ng, Zhichao Huang, and Yu Zhang. Low Resolution Image Categorization via Heterogeneous Domain Adaptation. **KBS’19 (CCF C)**.
- Yuan Yao, Yan Li, Yunming Ye, and Xutao Li. MLCE: An ensemble of decision cluster crotches for classification of multi-label data. **IJPRAI’20 (CCF C)**.
- Yuan Yao, Yan Li, Ke Wang, Zhichao Huang, and Yunming Ye. A Semi-supervised Clustering Method through Bottleneck Distance Exploration. **ICSS’16**.

开源项目

- **STN**: 基于柔性迁移网络的异构迁移学习方法。
- **CWAN**: 基于条件加权对抗网络的多源异构迁移学习方法。
- **DDA**: 基于判别式分布对齐框架的异构迁移学习方法。

受邀报告

极视角线上学术分享：迁移学习之异构域适应简介（[视频](#)）。

技能信息

- 英语：CET-6，且擅长专业文章写作
- 编程语言：Python、Tensorflow、Matlab、Hadoop、C/C++、Latex

获奖情况

- 哈尔滨工业大学优秀毕业生 2021
- 国家公派留学奖学金 2019-2020
- 哈尔滨工业大学一等奖学金 2015-2020
- “第九届服务科学国际会议” 最佳学生论文奖 2016
- 国家奖学金 2013, 2014
- 二零一三年高教社杯全国大学数学建模竞赛本科组二等奖 2013