**田佳杰**

男/28岁| 电话/微信：(+86)156-5221-1016 邮箱：[1051434278@qq.com](mailto:1051434278@qq.com)

**教育背景**

2013.09 - 2017.07 **中国石油大学(北京)** 数学专业 学士

2017.09 - 2020.07 **北京交通大学** 计算机科学与技术(推免) 硕士

**工作经历**

2021.06 – 现在 **蚂蚁集团有限公司(阿里巴巴)** 算法工程师

* 项目：tab3沉浸式召回
  + 项目介绍：在召回阶段，通过用户在服务和内容的历史行为序列，预测用户感兴趣的内容。
  + 个人职责：
    - 提出双塔模型，通过用户塔（用户画像、服务行为序列、内容行为序列）+内容塔（内容基础特征、内容理解特征），sample softmax loss优化用户对内容的点击。
    - 负样本：用户曝光未点击+随机
  + 工作成果：
    - 双塔模型acc=70%，auc=94%，recall@100=10%。
    - 正在线上做实验。
* 项目：服务与内容的相关性置信度
  + 项目介绍：通过分析服务与内容之间的相关性，利用用户在服务的访问行为预测用户的内容的偏好。
  + 个人职责：
    - 将服务与内容的相关性划分为四档：明确强相关、偏强相关、弱相关、不相关。
    - 通过SimCSE对比模型，学习服务与内容的相关性。
  + 工作成果：
    - 整体模型F1=75%+，其中“明确强相关”、“不相关”准确率90%。
* 项目：服务周期性标签
  + 项目介绍：通过对服务的短期、长期复访行为分析，提高首页纵横版高频服务的曝光比例，提升服务的pvctr。
  + 个人职责：
    - 针对服务的短期复访行为，提出对“一天内多次访问”的服务提权，降低疲劳度打压。
    - 针对服务的长期复访行为，提出对“访问周期天数”的服务提权，降低疲劳度打压。
  + 工作成果：
    - 标签“一天内多次访问”推全：命中服务的点击pv增加33%。
    - 标签“访问周期天数”推全：命中服务的点击pv增加23%。
* 项目：门店结构化
  + 项目介绍：对餐饮等行业的服务分析门店名称和门店地址，应用于搜索场景下的门店下挂，将小程序的商家信息提供给用户做决策，缩短用户路径，提高用户转化率。
  + 个人职责：
    - 通过分析餐饮行业不同小程序的门店名称的所在页面，提出“页面类型判断-门店名称提取”链路。
    - 通过多模态模型mmbt识别不同的页面类型：点单页、门店页。
    - 通过页面结构化模型block-bert抽取门店名称。
  + 工作成果：
    - 页面类型模型准确率95%+
    - 门店名称模型准确率95%+
* 项目：服务范围推断
  + 项目介绍：对服务的适用服务范围进行合理推断，应用在搜索等场景，提高推搜准确性。
  + 个人职责：
    - 通过不同信息（页面文本、用户行为、模板小程序等）进行服务内容推断，并根据一定的策略进行合并汇总。
    - 用NER建模，从页面文本抽取门店名称，并用高德返回门店地址，利用Sentence-pair Classification方法对高德返回结果审核
    - 通过聚合账单、足迹、全域的用户行为，推断出服务的常用地理范围
    - 通过对小程序模板分析数据，解析采集解码地理信息
  + 工作成果：
    - 文本抽取门店名称的F1=92%，高德门店名称匹配F1=90%。
    - 整体服务范围覆盖率 95%+，准确率85+。
* 项目：主动学习
  + 项目介绍：业务模型通常需要大量的人工标注数据进行训练，希望在数据标注的过程中，通过主动学习的方法，同时保证模型的效果和降低标注成本。
  + 个人职责：
    - 验证“基于未标注样本池+基于不确定采样”的主动学习方法在不同任务类型(多分类、层次分类等)的效果。
    - 在小样本场景下，提出利用度量学习预训练模型，结合主动学习的置信度区间+不确定区间，快速提升模型性能，召回正样本。
  + 工作成果：
    - 主动学习在不同任务类型中，模型效果下降约0.3%，save\_rate=50%，并成功应用于下游业务。
    - 小样本场景下，召回率=94.7%的情况下，save\_rate=96%。

2020.08 - 2021.04  **云从科技有限公司**  图像算法工程师

* 项目：视频浓缩
* 项目介绍：对于长时间稀疏的视频，通过视频浓缩系统，提取行人车辆轨迹，提高视频浏览速度，快速定位行人车辆信息。
* 个人职责：
  + 负责视频浓缩系统中前景分割、目标跟踪、轨迹浓缩、视频输出、摘要加速浓缩等模块的构建与优化。
  + 负责在线视频、离线视频的摘要与浓缩的pipeline的开发工作。
* 工作成果：
  + 训练前景分割模型BiseNetV2，IOU 92%。
  + 已经完成在线/离线视频的摘要/浓缩的pipeline的开发：双卡视频浓缩可达450fps，可支持18路视频并发；单卡视频浓缩可达350fps，支持14路视频并发。

2019.07 – 2019.09 **旷视科技** 实习生

* 项目：大库近似最近邻检索
  + 项目介绍：根据人脸特征，在大库中快速检索出最相似的K张图片。
  + 个人职责：
    - 对比PCA、DeepMDS(CVPR2019)等特征降维方法，学习OPQ等向量量化方法。
  + 工作成果：
    - 特征降维方面，PCA效果更好，OPQ可以实现更快检索。

2019.05 – 2019.11 **北京交通大学** 硕士研究生

* 项目：行人重识别项目UMDA
  + 项目介绍：使行人重识别项目能自适应多个无标签目标域
  + 个人职责：
    - 提出利用StarGAN同时对齐源域和多个目标域的分布
    - 提出利用camera id 区分不同的目标域
  + 工作成果：
    - CUHK03—> Duke, Market-1501的性能(top-1: 50.1%, 66.6%)总体优于其他方法

2018.10 – 2019.03 **北京交通大学** 硕士研究生

* 项目：行人重识别项目UDA
  + 项目介绍：使行人重识别项目能迁移无标签目标域
  + 个人职责：
    - 提出ImitateModel补充无标签数据集的absolute label 和 pairwise label
    - 提出分类损失约束 absolute label，学习目标数据集的整体特征分布
    - 提出三元组损失约束 pairwise label，学习目标数据集的 camera-inariance
  + 工作成果：
    - 在 Market-1501数据集，rank-1 达到72.4%，相对提高10.2%，mAP 达到 40.1%，相对提高8.7%

**论文**

* **一作**
* Jiajie Tian, Qihao Tang, Rui Li, Zhu Teng, Baopeng Zhang, and Jianping Fan. 2021. A Camera Identity-guided Distribution Consistency Method for Unsupervised Multi-target Domain Person Re-identification. ACM Trans. Intell. Syst. Technol. 12, 4, Article 38 (August 2021), 18 pages. https://doi.org/10.1145/3454130
* Tian, Jiajie et al. “Imitating targets from all sides: an unsupervised transfer learning method for person re-identification.” International Journal of Machine Learning and Cybernetics 12 (2019): 2281 - 2295.

**获奖**

|  |  |
| --- | --- |
| * 研究生一等学业奖学金 | * 数学建模国际赛二等奖 |
| * 本科生一等奖学金 | * 本科生国家励志奖学金 |

**自我评价**

* 座右铭：也许一个人也能走得很快，但是一群人才能走得更远