

马浩杰

17342551730 · HaojieMa@outlook.com · angelhunt ·

个人信息

男，1994 年出生

求职意向：算法工程师

工作经验：约1年

教育经历

- 硕士，南京大学，计算机科学与技术专业，2017.9~2020.7
- 学士，南京大学，计算机科学与技术专业，2013.9~2017.7

工作经历

- 微软中国公司，m365 部门，2020.6~至今
大数据平台用户/集群异常检测
基于scope语法树静态分析实现的data lineage系统

科研经历

- 两篇CCF A类一作会议论文
- Unsupervised Human Activity Representation Learning with Multi-task Deep Clustering (**UbiComp 2020**)
 - 构建了多任务无监督深度聚类，用于人类活动识别
 - 模型通过对聚类，分类和自编码三个任务的联合学习，获取了高效的隐层表征
 - 模型在无标签的情况下能超过主流无监督方法，并且通过少量标签的fine-tuning可以逼近监督算法性能
- AttnSense: Multi-level Attention Mechanism For Multimodal Human Activity Recognition (**Ijcai 2019**)
 - 多模态场景下人类的行为识别
 - 通过self attention模型进行了多模态传感器信号的dynamic fusion

项目经历

- Humor Generation and Classification (幽默生成与分类)
 - 使用GPT2-chinese并通过humor dataset进行fine-tuning, 用于生成中文笑话
 - 使用BERT与Fine-Tuned的分类器在生成的笑话数据集进行训练，用于分类
- 大数据集群异常检测
 - 基于PCA的用户行为异常检测与拉黑
 - 集群节点异常检测，例如任务阻塞，job失败率过高
- 基于Intel MKL加速spark MLlib算法性能 (本科毕设)
 - 通过MKL实现了PCA, ALS, Logistic Regression等算法
- NJU-ProjectN
 - 实现了完整的流水线CPU, c语言子集编译器, MIPS模拟器和简易操作系统
- 造的一些轮子
 - python实现的机器学习算法，例如决策树，逻辑回归，线性回归算法
 - 基于attention网络的文档分类模型 (复现HAN)
 - 使用hadoop map reduce构建了多标签聚类算法，用于分析金庸小说，展示小说人物关系图
 - c++, MPI, cuda等库实现的并行机器学习算法

获奖经历

- 2015年-2017年 南京大学人民奖学金
- 2019年-2020年 国家奖学金