

作品说明书

多源异质大数据联合分析

项目类别 | 自然科学类论文

指导老师 | 张逸群、姬玉柱、骆文君

项目成员 | 张云帆、陈欣禧、谭泽熙、甘国基、龙志全、高逸菲、蔡升宏、陈俊仰

研究背景 充分开发数据资源，推进智能分析决策，成为重要命题

大数据 已成为推动社会进步和经济发展的核心资源



挑战性分析

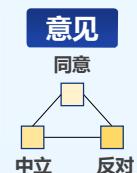
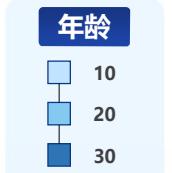
针对多源数据的异质和标签匮乏难题，创新联邦聚类系统性解决方案

信息融合损耗大

数据异质
对齐程度弱

定量数据

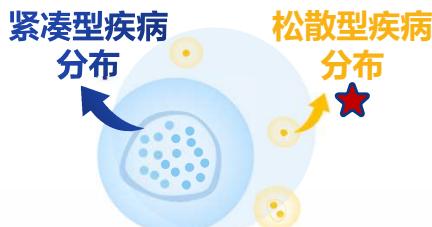
定性数据



异质特征数据距离
新度量

联邦聚类精度低

信息非完备
聚类依据匮乏



多源数据联邦聚类
新策略

知识更新时效差

数据动态
更新链路复杂

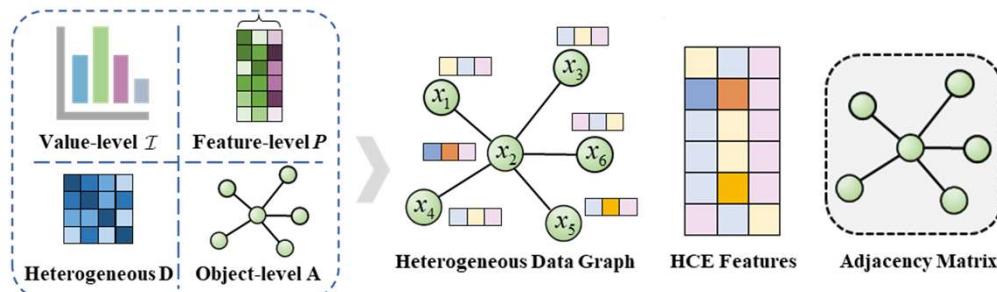


动态环境持续学习
新模型

创新点一 异质特征数据距离新度量 自研统一度量+多层次耦合编码

异质信息提取更充分，分析决策精度高，发表CCF-A类成果

层次耦合编码策略



信息融合损失

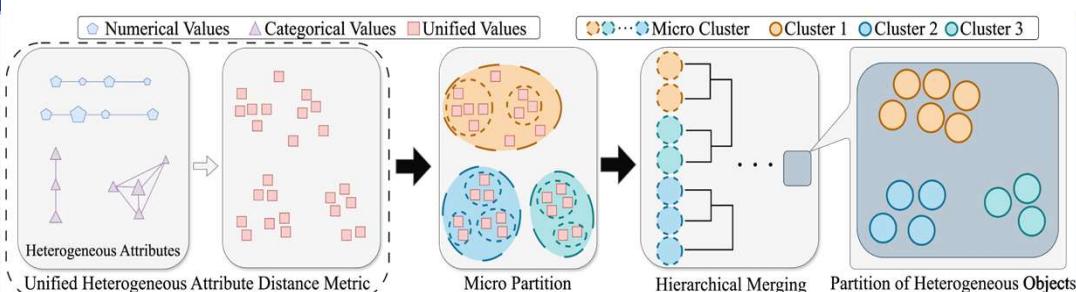
降低高达

21.02%

创新点二 多源数据联邦聚类新策略 多粒度层次脱敏+簇分布自组织

分布检测更细致，知识粒度自适应，发表SCI一区成果

微簇划分+分层合并策略



知识嗅探精度

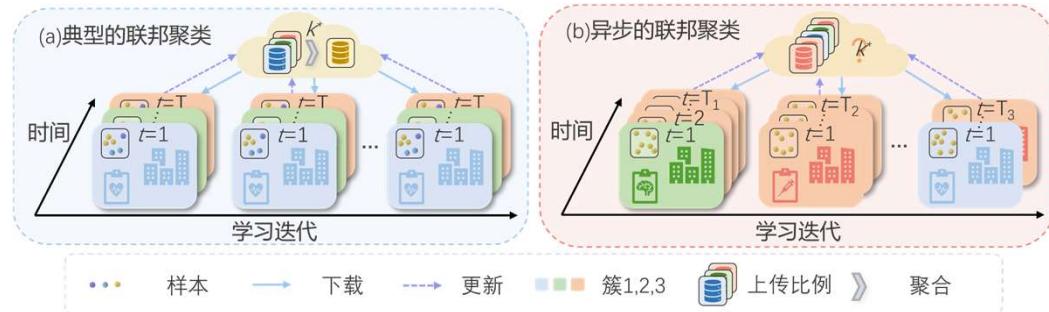
高达

90.05%

创新点三 动态环境持续学习新模型 异步更新平衡+漂移感知网络

分析鲁棒性强，决策精度再提升，发表CCF-A类成果

传统联邦vs异步联邦



知识更新时效

提升高达

74.28%

团队成员

张云帆 (负责人)

- 录用SCI一区期刊一篇 (一作)
- 录用CCF-A会议一篇 (一作)
- 录用CCF-C会议两篇 (一作)
- 录用IEEE会议一篇 (最佳论文奖)
- 在投SCI一区期刊一篇 (二作)
- 获得本科生国家奖学金
- 已被香港高校直博全奖录取

陈俊仰

- 录用SCI一区期刊一篇 (一作)
- 录用CCF-A会议一篇 (一作)
- 在投SCI一区期刊一篇 (一作)
- 保研至清华大学

蔡升宏

- 录用CCF-B会议一篇 (一作)
- 在投CCF-A会议一篇 (一作)
- 国家级大创项目第一负责人
- 多所港澳高校直博全奖条件录取

陈欣禧

- 录用SCI二区期刊一篇 (二作)
- 录用CCF-C会议一篇 (二作)
- 在投SCI一区期刊一篇 (二作, 指导老师一作)
- 国家级大创项目第一负责人

谭泽熙

- 录用SCI一区Top期刊一篇 (第一本科生作者)
- 录用SCI二区期刊一篇 (三作)
- 录用EI会议论文一篇 (一作)
- 国家级大创项目第一负责人
- 广东省级科技创新项目第一负责人

甘国基

- 在投SCI一区期刊一篇
- 在投CCF-B会议一篇
- 实审两项国家发明专利

龙志全

- 在投SCI一区期刊一篇
- 在投CCF-B会议一篇
- 实审两项国家发明专利

专家推荐

沈卫明

加拿大工程院院士
华中科技大学教授
福耀科技大学讲席教授

推荐理由 摘选

“ 对无监督联邦学习和异质数据分析的研究具有前沿性和前瞻性，项目所提出的方案为解决该领域的关键基础性技术难题提供了创新思路。 ”

蔡宏民

国家杰出青年科学基金获得者
华南理工大学教授

推荐理由 摘选

“ 成果构成了联邦异质数据分析方法体系，理论性和科学性较强，研究拓展了数据科学和机器学习研究领域。有望提升行业应用中的数据分析效能。 ”

团队已取得论文成果**12篇**, 其中**8篇**为本科一作
另有在投高水平论文**6篇**

SCI一区

2 篇



CCF-A类

2 篇



CCF-B类

1 篇



**IEEE
最佳论文奖**

1 篇

所获论文成果列表

No. 名称

1	Yiqun Zhang,..., Zexi Tan, Xiaopeng Luo, Yuzhu Ji*, Rong Zou Learning SelfGrowth Maps for Fast and Accurate Imbalanced Streaming Data Clustering, <i>TNNLS</i> , 2025. 探索动态数据分布自适应的高精度高效率非平衡数据聚类算法	SCI一区 IF: 8.9
2	Junyang Chen, Guoheng Huang,..., and Zhixin Huang. Quaternion Cross-Modality Spatial Learning for Multi-Modal Medical Image Segmentation, <i>JBHI</i> , 2024. 面向医学影像分割应用的距离度量空间学习深度模型	SCI一区 IF: 6.8
3	Yunfan Zhang, Rong Zou, Yiqun Zhang*, ..., Kangshun Li Adaptive Micro Partition and Hierarchical Merging for Accurate Mixed Data Clustering, <i>CAIS</i> , 2024. 提出数据微簇划分机制，引导高精度异质数据聚类	SIC二区 IF: 4.6
4	Chuyao Zhang, Xinxi Chen, Zexi Tan,..., Yuzhu Ji, Yiqun Zhang* ANDI: ANy-type Attributed Data Imputation via Cluster-Guided Missing Value Inference, <i>CIS'24</i> . 聚类引导的异质特征数据缺失值推断与补全	SIC二区 IF: 2.3
5	Yunfan Zhang, Yiqun Zhang*, Yang Lu, Mengke Li, Yiu-ming Cheung Asynchronous Federated Clustering with Unknown Number of Clusters, <i>AAAI'25</i> . 本地节点异步传输且合理簇数未知情形下的联邦聚类范式	CCF-A
6	Junyang Chen, Yuzhu Ji , Rong Zou, Yiqun Zhang, Yiu-ming Cheung QGRL: Quaternion Graph Representation Learning for Heterogeneous Feature Data Clustering, <i>SIGKDD'24</i> . 引入超复空间深度表征学习机制突破异质特征数据聚类精度瓶颈	CCF-A
7	Shenghong Cai, Yiqun Zhang*, ..., Hong Jia, Peng Liu Robust Categorical Data Clustering Guided by Multi-Granular Competitive Learning, <i>ICDCS'24</i> . 拓展多粒度竞争学习理论提升聚类分析鲁棒性	CCF-B
8	Yunfan Zhang, ..., Yiqun Zhang*, Yiu-ming Cheung Towards Unbiased Minimal Cluster Analysis of Categorical-and-Numerical Attribute Data, <i>ICPR'24</i> . 提出异质特征数据最小簇理论实现低归纳偏执聚类分析	CCF-C
9	Pengkai Wang†, Yunfan Zhang†, Yiqun Zhang*, ... Yiu-ming Cheung Clustering by Learning the Ordinal Relationships of Qualitative Attribute Values, <i>IJCNN'24</i> . 进行异质特征数据取值顺序关系学习以增强聚类精度	CCF-C
10	Haoyi Xiao, Xinxi Chen, Xiaopeng Luo*, ..., Wei Ai MACL: Metric and Attribute Space Co-Learning for Qualitative Data Clustering , <i>ICIC'25</i> . 提出度量空间和属性子空间协同学习机制增强聚类精度	CCF-C
11	Rong Zou, Yunfan Zhang, Yiqun Zhang,..., Yiu-ming Cheung* Federated Clustering with Unknown Number of Clusters, <i>IEEE DOCS'24</i> . 针对普遍的真实簇数未知情形提出簇分布自探索联邦聚类算法	EI IEEE 最佳论文奖
12	Chuyao Zhang, Xinxi Chen, Zexi Tan,..., Yuzhu Ji, Yiqun Zhang* ANDI: ANy-type Attributed Data Imputation via Cluster-Guided Missing Value Inference, <i>CIS'24</i> . 聚类引导的异质特征数据缺失值推断与补全	EI

注： 团队本科生成员 指导老师 通讯作者*

查新结果

项目名称

多源异质大数据联合分析

查新机构

中国科学院上海科技查新咨询中心
(国家一级)

查新结论

国内外文献检索中
未见相同文献报道

本查新项目具有新颖性



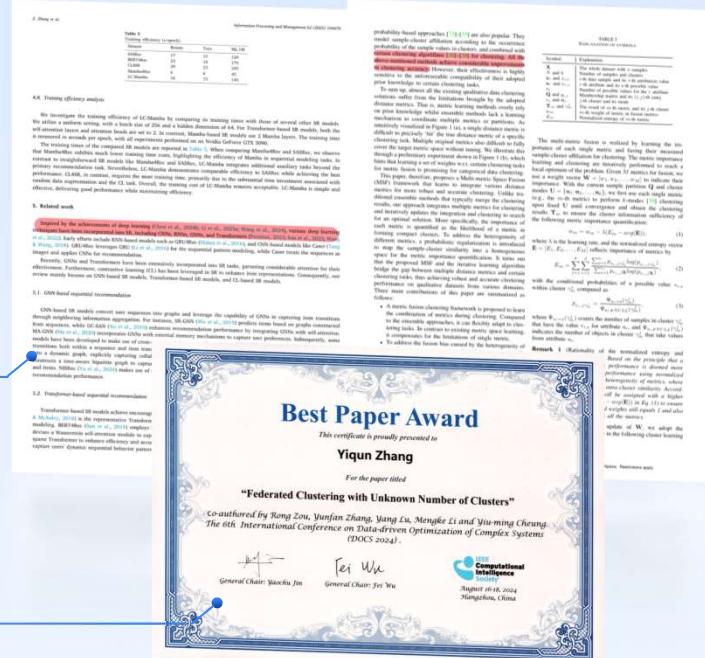
学术认可

论文成果获知名学者

- 香港浸会大学讲席教授张晓明
- 电子科技大学教授杨波

多次正面引用与评价

论文成果获 IEEE最佳论文奖 1项



现实应用价值

- ✓ 与各三甲医院展开广泛合作与应用部署
(如广州医科大学附属第三医院、深圳市妇幼保健院)
- ✓ 成果价值广受合作单位认可



广泛应用拓展

