

周杰辉

Github: github.com/algzjh Email: jiehuizhou1997@gmail.com
个人主页: algzjh.github.io 手机/微信: 18932449056

我现于浙江大学计算机学院，CAD&CG 国家重点实验室攻读博士学位，导师是陈为教授。我的科研主题是医疗健康异构数据挖掘与可视分析，具体涉及 可视分析 (visual analytics)，因果分析 (causal analysis)，异常检测 (out-of-distribution detection) 以及隐私保护 (privacy preservation) 等领域。我正在寻找一份大数据分析方向的工作，包括以下领域：数据可视分析以及可信人工智能，如可解释 AI 和因果推理等。

教育经历

- 中南大学 长沙，中国

Bachelor of Computer Science & Technology; GPA: 91.53/100 *Sep 2015 - June 2019*

课程：数据结构，算法分析与设计，操作系统，数据库，计算机网络，概率论与数理统计，机器学习等。
- 浙江大学 杭州，中国

PhD of Design Science; State Key Lab of CAD&CG; GPA: 88.49/100 *Sep 2019 - June 2024 (plan)*

课程：大数据可视化前沿技术，大数据存储与处理，三维动画与交互技术，云计算原理与应用开发等。

科研工程 & 论文发表

- 医保数据时空关联分析系统（多维关联分析，时空数据可视化）**：应用落地导向。问题：多种类型的医保欺诈行为在参与主体和时空等维度上关联交织，分析难度大。方法：(1) 设计频繁模式挖掘和异常检测模型快速筛查多类可疑行为，如多地开药和虚假住院等；(2) 开发多视图联动的可视分析系统，提供时序和地理可视化，支持用户交互式地挖掘、探索和验证关联的欺诈行为。评估：系统实际部署于杭州市富阳区医保局，帮助检测出涉及多类欺诈行为的参保人和医生，获得用户的积极评价。
Publication: Zhou Jiehui, Zhu R, et al. MedicareVis: a Joint Visual Analytics Approach for Anti-Fraud in Medical Insurance. *Journal of CAD & CG*, 2021. (EI) [\[Paper\]](#)
- 医疗共谋欺诈行为可视分析（多源分布式数据，图网络建模，异常检测，人机协同学习）**：应用落地导向。问题：医保共谋欺诈涉及多人多地作案，且难以与正常患者行为（如慢性病群体的定期复诊）相区分，导致检测结果假阳性率高，需要大量时间人工校验。方法：(1) 结合共谋欺诈的时空关联特征，构建加权图网络刻画欺诈者之间的共同就诊行为，再利用社团检测模型挖掘可疑的共谋欺诈群体。(2) 提出基于疾病和药品的患者相似性度量，帮助用户判断患者多次就诊的合理性。(3) 开发交互式系统，设计图和事件序列等可视化形式，支持用户高效地比较、探索和验证可疑的共谋欺诈行为。评估：基于富阳地区 2020-2021 年的数据，系统共发现共谋欺诈患者 435 人，涉及机构数 66 家，医疗总金额达 50 多万元。
Publication: Zhou Jiehui, Wang X, et al. FraudAuditor: A Visual Analytics Approach for Collusive Fraud in Health Insurance. *IEEE Transactions on Visualization and Computer Graphics*, 2023. (CCF-A) [\[Paper\]](#) [\[Slide\]](#) [\[Demo\]](#)
- 可视化图表的隐私保护发布方法（数据跨域安全，差分隐私，贝叶斯网络，人机交互）**：科研导向。问题：可视化图表的发布面临着隐私泄露和传播洞察的权衡。方法：(1) 提出模式约束来建模用户对于不同数据模式的偏好，并用其引导基于贝叶斯网络的差分隐私保护的数据生成模型；(2) 开发可视化系统支持用户交互式地选择模式并指定其重要性，理解模型的隐私保护过程，以及比较不同的隐私保护方案。评估：基于聚类、相关性和排名等模式效用指标的定量实验证明，在中等保护级别下，所提方法显著性优于现有隐私保护的数据生成模型基线（显著性水平 p 值 <0.01 ，平均效应量 Cohen's $d=0.5$ ）。
Publication: Zhou Jiehui, Wang X, et al. DPVisCreator: Incorporating Pattern Constraints to Privacy-preserving Visualizations via Differential Privacy. *IEEE Visualization Conference*, 2022. (CCF-A) [\[Paper\]](#) [\[Code\]](#) [\[Demo\]](#)
- 因果异构性可视分析（子集分布异构，可解释聚类，因果推断，人机交互）**：科研导向。问题：在数据分布异构的场景下，原因难以被预先定义，这迫使数据分析人员反复尝试原因假设、亚组分析和效应验证。方法：(1) 引入概念来描述具有相似的特定属性值的数据单元，以提高可解释性和灵活性；(2) 提出了基于概念的模型 HetMatch，通过匹配决策规则来快速识别具有潜在有效因果关系的异构性上下文；(3) 开发具有联动可视化视图的系统，支持用户交互式地编辑概念，比较子群之间的因果关系，并根据匹配的单元解释因果效应。评估：基于模拟数据的定量实验证明了 HetMatch 在超过 10 个变量的情况下效应估计误差小于基线方法 CausalTree。医疗诊断和手机故障分析等实际场景的案例分析证明了系统的可用性。
Publication: Zhou Jiehui, Wang X, et al. CausalPrism: A Visual Analytics Approach for Concept-based Causal Heterogeneity Exploration. (under review)

实习经历

- 阿里巴巴 线下实习

Engineering Intern *June 2020 - Aug 2020*

- **项目:** 淘宝闲鱼用户社交数据的关系挖掘与分析, 辅助好友和商品推荐。
- **关键词:** 分布式计算, 社交网络分析。

之江实验室

线下实习

Research Intern

Sep 2020 - May 2021

- **项目:** 海量动态图数据可视分析系统, 支持分析不同节点之间的时序关联和演化模式。
- **关键词:** 数据融合, 动态图, 可视分析系统。

OPPO

线下实习

Research Intern

Sep 2022 - Jan 2023

- **(1) 项目:** 基于手机埋点数据的故障根因分析, 挖掘监控变量之间的因果关系和效应, 开发系统探索和验证因果影响路径。
- **(1) 关键词:** 因果发现, 根因路径挖掘, 因果可解释性。
- **(2) 项目:** 评价机器学习模型在不同使用场景下的表现。利用自监督学习埋点表格数据的表征, 再基于用户反馈标注少量数据, 采用半监督微调得到评分预测模型, 最后利用 SHAP 等可解释性方法得到各个特征对模型评分的贡献。
- **(2) 关键词:** 自监督学习, 半监督学习, 多维表格数据, 模型可解释方法。

基础技能

- **语言:** Python, JavaScript, C++, Pytorch, Scikit-learn, Django, React
- **工具:** Git, MySQL, Docker, Linux Server
- **软性竞争力:** Writing, Project Management, Time Management, Public Speaking
- **社区参与:** 知乎 (可信人工智能), 语雀 (因果推理, 找论文技巧, 论文读写经验)

荣誉与奖项

- 国家奖学金 (2016)
- ACM-ICPC 亚洲区域赛南京站, 徐州站银奖 (2018)
- 三好研究生 (2021)