



## 教育经历

### 南京信息工程大学-应用统计学（本科）

全国大学生英语竞赛一等奖，CET6(597)

江苏南京

2017.09-2021.06

### 浙江工商大学-理学统计学（硕士）

一作发表深度生成模型论文一篇(CCF-A)，图推荐系统论文一篇(中科院一区&CCF-B)  
和深度强化学习论文一篇(中科院二区)，另有两篇以一作/共同一作投稿至TKDE及ICML24。

浙江杭州

2021.09-2024.01

## 论文发表

- Xu C, Wang H, Wang W, et al. Geometric-Facilitated Denoising Diffusion Model for 3D Molecule Generation. Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 2024.
- Han X, Shan C, Shen Y, Xu C, et al. Training-free Multi-objective Diffusion Model for 3D Molecule Generation. Proceedings of the Twelfth International Conference on Learning Representations (ICLR), 2024.
- Xu C, Zhang Y, Chen H, et al. A fairness-aware graph contrastive learning recommender framework for social tagging systems. Information Sciences, 2023, 640: 119064.
- Xu C, Zhang Y, Wang W, et al. Pursuit and evasion strategy of a differential game based on deep reinforcement learning. Frontiers in Bioengineering and Biotechnology, 2022, 10: 827408.

## 实习经历

### 之江实验室-图计算研究中心

智能计算计算医药（大规模图预训练，分子性质预测，Diffusion-based分子生成）

图算法工程师

2022.07-2023.08

- 创造性提出基于全局注意力的E(n)等变图网络，结合原子自身性质与分子间相互作用，实现时下最优的3D分子学习性能。同时将此图网络与扩散模型（Diffusion）结合，实现时下最优的3D分子生成性能，相关文章发表于AAAI-24；
- 负责Diffusion-based生成模型研究，撰写Diffusion-based图生成领域综述，相关文章投稿至TKDE；
- 负责OGB-LSC NeurIPS22竞赛，构建模型实现超大规模分子图性质学习，取得全球第11名；
- 复现、部署大规模图预训练模型及分子构象生成模型。

## 项目经历

### 真实图数据的分布外泛化

第一作者

论文(投稿至ICML-24)

2023.10-2024.02

- 针对图OOD泛化领域的方法在手工数据集上表现良好但在真实数据中表现不佳的现状，提出更符合真实场景的图OOD泛化算法；
- 不同于先前因果推断方法强调挖掘变量间的因果关系或独立性关系，本文从真实数据出发，提出发掘变量间的条件依赖性；
- 根据变量间的相互条件依赖关系，构建变分推断的框架DEROG，结合EM算法对提出的DEROG框架进行优化；
- 提出更符合真实场景的环境对齐函数和图主干的对比损失，提升DEROG对隐变量的推理能力；
- 文章“Improving Graph Out-of-distribution Generalization on Real-world Data”投稿至机器学习顶会（CCF-A）ICML-24。

### 异构图预训练上的对抗攻击方法

第二作者

论文(投稿至KDD-24)

2023.10-2024.02

- 目前大量的异构图预训练模型被提出，但缺少针对其鲁棒性的研究，本文首次提出了针对异构图预训练的无监督攻击模型HGAC；
- HGAC考虑了攻击的隐蔽性和合理性，采用边缘重连的攻击策略，并使用节点聚类作为伪标签来指导攻击的优化；
- 在节点分类、节点聚类、链接预测、可视化四种下游任务都证明了HGAC良好的攻击性能；
- 文章“Unsupervised Heterogeneous Graph Rewriting Attack via Node Clustering”投稿至数据挖掘顶会（CCF-A）KDD-24。

### 几何辅助的3D分子图扩散生成

第一作者

论文

2022.11-2023.08

- 基于扩散生成模型，构建去噪扩散概率模型GFMDiff，生成创新且有效的3D分子结构；
- 将扩散模型自身特点与原子间相互作用力本质相结合，构建出既能学习不同原子化学特性，又能学习任意原子间的相互作用强弱的全局注意力机制，构建创新的等变图网络作为去噪过程的内核；
- 针对扩散模型在离散图数据中表现不佳的问题，创新性提出GFLoss，通过几何信息和理化性质引导模型生成真实且合理的图结构；
- 文章“Geometric-facilitated denoising diffusion model for 3D molecule generation”已被人工智能顶会（CCF-A）AAAI-24接收。

## 免训练的基于条件的分子生成

作者

论文

2023.09-2023.10

- 针对现有Diffusion框架在基于条件的生成任务中表现出的低效，不灵活和不佳表现，提出免训练的条件3D分子生成模型MuDM；
- 提出对条件函数梯度进行MC采样，并最大化最终预测结果的条件期望分布，解决了生成模型面对不同条件时重复训练的问题；
- 在极大提升训练效率的同时，MuDM在所有基准测试中结果显现出数量级提升；
- 文章“Training-free Multi-objective Diffusion Model for 3D Molecule Generation”被机器学习顶会 ICLR-24 接收。

## 基于扩散的图生成模型综述

共同第一作者 & 实际撰稿人

论文(投稿至TKDE)

2023.02-2023.08

- 系统性总结时下最先进的扩散生成模型在图生成任务上的应用，对该领域研究做了时下最全面的总结与讨论；
- 系统性阐述主流扩散模型范式，讨论现有Diffusion模型在适配图生成时的不同方法和挑战；
- 阐述并比较扩散图生成在计算化学，动作生成等领域的应用；
- 讨论扩散模型在图生成任务中面临的不足与挑战，并对未来的研究方向提出展望。
- 文章“Diffusion-based graph generative methods”投稿至数据挖掘顶刊 (CCF-A) IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING (TKDE)。

## OGB-LSC NeurIPS 2022

负责人

竞赛

2022.08-2022.11

- 对超大规模分子数据(超300W)进行表征学习，预测目标分子能量隙；
- 构建HFAGNN模型，运用Hybrid模块，学习原子自身化学性质，同时运用Bessel方程提取2元组与3元组间的原子3D信息，将分子2D拓扑信息与3D结构信息结合，在满足三维空间等变性的同时实现对分子性质的高效学习；
- 使用多卡并行实现高效运算，在较小规模参数量的情况下取得榜单第11名。

## 基于图对比学习的公平的标签感知推荐系统

第一作者

论文

2022.05-2022.12

- 基于GNN和对比学习范式，构建TAGCL模型进一步提升基于标签感知的个性化推荐性能，同时显著提升推荐结果的公平性；
- 分别构建<用户-标签>和<标签-商品>的子图，通过添加经过softmax归一化的随机扰动进行表征增强，用于建立对比学习任务；
- 分别基于交叉熵损失，基于Tag负采样的二部图损失和改进知识图谱算法TransT，构建目标函数，分别用于推动生成更公平的推荐和提升二部图之间的一致性；
- 实验显示模型推荐精度较标签感知推荐领域的SOTA模型至少提升5%，同时在歧视性较强的数据集上显著提升结果的公平性；
- 文章“A fairness-aware graph contrastive learning recommender framework for social tagging systems”发表于期刊(中科院一区(TOP) & CCF-B) INFORMATION SCIENCES。

## 电商用户商品价值评估与基于图神经网络的个性化推荐系统

组长

全国统计专业研究生案例大赛

2022.02-2022.03

- 基于真实电商数据，实现商品从盈利能力、畅销水平和退货率等方面的价值评估，同时基于价值对用户群体进行有效划分，最后对用户进行个性化推荐。
- 通过PCA降维与K-Means++聚类，实现对商品从盈利能力，畅销水平和退货率等多个维度的有效划分。
- 运用改进RFM模型，对用户特征做特征交叉，实现对用户从消费能力，下单频率和退货水平等多维度的有效划分。
- 基于LightGCN模型，创新性的提出了WideGCN模型，实现对基于商品和用户的16维特征这一旁信息的有效利用。推荐性能较SOTA模型有2%到5%的提升。
- 该项目获全国三等奖

## 空间抗干扰UWB定位算法

组长

全国研究生数学建模竞赛

2021.10

- 构建方程实现分别基于3锚点和4锚点的空间UWB定位算法，实现较好的抗干扰性能。
- 利用SVM实现测量数据受干扰情况的准确判别。
- 利用上述定位算法与卡尔曼滤波实现物体在三维空间内的轨迹追踪。
- 获得第18届华为杯全国研究生数学建模竞赛三等奖。

技能

技能 Python, PyTorch, PyG, DGL, Tensorflow, Git, Linux, SQL,  $\LaTeX$   
外语 CET4 (597) , CET6 (597)

荣誉奖励

2018.05	2018年全国大学生英语竞赛 (NECCS)	全国一等奖
2019.05	南京信息工程大学第十三届数学建模竞赛	三等奖
2021.10	华为杯第十八届全国研究生数学建模竞赛	全国三等奖
2022.08	第五届全国统计专业研究生案例大赛	全国三等奖
2023.06	泰迪杯数据挖掘竞赛	三等奖
2022.11	浙江工商大学研究生学业奖学金	一等奖
2023.11	浙江工商大学研究生学业奖学金	一等奖
2023.11	国家奖学金	获奖