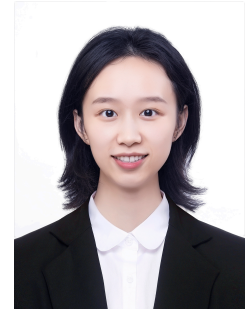


# 王宇星

✉ wang-yuxing@foxmail.com

☎ 15927541048

🏠 中山大学，广州



## 教育经历

- 2021 至今    ■ 博士    中山大学（985），计算机科学与技术，知识图谱数据挖掘方向
- 2018 – 2021    ■ 硕士    华中农业大学（211），生物信息学，生物自然语言处理方向
- 2014 – 2018    ■ 本科    华中农业大学（211），生物信息学

## 科研经历

### 知识图谱数据挖掘方向

- 2023.10 至今    ■ 图神经网络的可解释性研究。[人工智能算法设计]  
该项目针对神经网络的黑匣子问题提出基于原型的神经网络算法，让神经网络在预测知识图谱中关系的同时可视化其相关子图，使得其决策透明可见。
- 2022.10-2023.10    ■ 基于因果学习的神经网络模型设计。[神经网络模型搭建]  
该项目构建大规模知识图谱，开发了一个基于动态超图和因果学习的神经网络模型，动态超图能够灵活获取大型知识图谱数据中的复杂关系，而将因果学习与神经网络方法结合，能够找到大型知识图谱数据中包含因果关系的部分，并利用这一部分进行数据关系预测，为预测结果提供更详细的解释信息。
- 2021.9-2022.10    ■ 基于信息传递路径的图神经网络方法改进。[人工智能算法改进]  
该项目是对人工智能领域图神经网络的方法改进，从路径的角度挖掘知识图谱网络结点之间的关系，可应用于大数据网络中的关系预测问题，并从数据的角度提供预测结果的路径证据。

### 自然语言处理方向

- 2020.3-2021.4    ■ 基于异质数据融合的算法设计。[算法主要设计人之一]  
本项目通过集合多方异质生物数据，构建大规模异构网络，设计张量分解算法构建模型，使得模型学习到全面的生物数据信息，在基因疾病关联预测应用中取得较好效果。
- 2019.8-2020.1    ■ 利用语料库进行文本知识挖掘。[项目流程设计]  
基于大语言模型 *BERT*，通过挖掘生物医学文本中的突变事件，用于疾病药物预测。

## 科研经历 (continued)

- 2019.6-7/2020.1-2 ■ BioNLP-OST2019/2020 AGAC 任务。[国际赛事组织]  
作为任务发起者参与了 2019 年及 2020 年生物自然语言处理公开分享任务，发布并评审利用语料库进行生物医学文献知识挖掘任务，其中 2020 年为针对新型冠状病毒的文献知识挖掘任务。
- 2018.1-12 ■ 活跃基因注释语料库构建。[语料库设计]  
AGAC 语料库标注生物医学文本中的基因突变事件，即基因突变导致后续生物过程变化最后与某一疾病相关的信息。项目中，本人负责语料库的标签关系设计和标注规则设定。

## 邀请报告

- 2023.04 ■ 27th Annual International Conference on Research in Computational Molecular Biology, Istanbul, Turkey.  
报告题目: “Gene based message passing for drug repurposing[基于基因信息传递路径的药物重定位]”
- 2020.2 ■ BioNLP Open Shared Task 2020 Workshop, Japan(online).  
报告主题: “Task Review of AGAC Track in BioNLP OST 2020”
- 2019.11 ■ BioNLP Open Shared Task 2019 Workshop, Hongkong, China.  
报告主题: “Task Review of AGAC Track in BioNLP OST 2019”

## 发表文章

- 1 [第一作者/CCF-B 类会议送审中]. (2024). Cdhl: Drug repurposing through causal dynamic hypergraph learning discovery. *BIBM 2024*.
- 2 [第一作者/Cell 子刊]. (2023). Gene based message passing for drug repurposing. *Isience*, 26(9).
- 3 [共一作者/SCI 一区 Top 期刊]. (2021). Bridging heterogeneous mutation data to enhance disease gene discovery. *Briefings in Bioinformatics*, 22(5), bbab079.
- 4 [第一作者/BioNLP 赛事]. (2019). An overview of the active gene annotation corpus and the bionlp ost 2019 agac track tasks. *Proceedings of The 5th Workshop on BioNLP Open Shared Tasks*, 62–71.
- 5 [第一作者/CCF-B 类会议]. (2019). An active gene annotation corpus and its application on anti-epilepsy drug discovery. *BIBM 2019, San Diego, U.S*, 512–519.
- 6 [第二作者/SCI 期刊]. (2019). Gof/lof knowledge inference with tensor decomposition in support of high order link discovery for gene, mutation and disease. *Math Biosci Eng*, 16(3), 1376–1391.

## 获得荣誉

- 硕士阶段 ■ 国家奖学金 (2020)，三好研究生 (2018-2020)，甲等学业奖学金 (2018-2020)，优秀研究生毕业论文，优秀研究生毕业生 (2021)
- 博士阶段 ■ 二等博士研究生学业奖学金 (2022)，三等博士研究生学业奖学金 (2021,2023)