

# 林学渊

188-1157-2707 | linxy59@mail2.sysu.edu.cn | 北京海淀  
https://github.com/LinXueyuanStdio  
知识图谱

## 教育经历

北京邮电大学	2020年09月 - 2023年06月
计算机科学与技术 硕士 计算机学院 ( 国家示范性软件学院 )	北京
● 荣誉/奖项：一等学业奖学金 ( 2020-2021 )	
中山大学	2016年09月 - 2020年06月
数学 本科	珠海
● 专业课GPA：4.2 / 5.0 ( 专业前10% )	
● 荣誉/奖项：“发现杯” APP 开发组全国一等奖 ( 2017 )、国家助学金 ( 2016-2017 )	

## 研究经历

高血压知识图谱诊疗辅助系统	2021年01月 - 至今
算法研发	北京
为了提高基层医院的高血压诊疗能力，辅助普通医生达到三甲医院医生的水平，本研究提出基于高血压知识图谱的诊疗辅助系统。	
1. 参与高血压认知图谱的构建，基于权威医学资料，构建患者画像、高血压知识超图等，包含 136 个三元组、311 条推理规则	
2. 负责逻辑推理模块，基于患者当前身体各项指标，按权威医学规则进行模糊化逻辑推理，给出高血压风险分层、治疗方案、药物决策。目前推理相关技术已投稿会议 NIPS2022 ( CCF A, 两篇 )，专利申请中	

## 项目经历

“千企千面” 政策企业匹配	2021年09月 - 2022年03月
算法研发	北京
为了方便企业从海量政策中及时获取惠企信息，加快政府对企业的扶持，本项目给企业推荐惠企政策	
1. 负责政策画像的构建，利用新词发现算法抽取 8k+ 具有政策特征的词，利用聚类算法辅助从 12 个维度构建政策画像	
2. 负责政策、企业画像的自动化打标，利用预训练语言模型进行打标，准确率约 80%	
3. 负责政策新闻分类算法，综合利用人工规则和预训练语言模型，准确率约 90%，节省用于筛选政策的人力到 1 人/周	
4. 搭建了自动化流水线，打通爬虫数据->清洗->分类初筛->画像打标->政策推荐全流程	

## 实习经历

阿里巴巴达摩院 ( 杭州 ) 科技有限公司	2022年06月 - 至今
研究型实习生	
负责研究面向复杂逻辑推理的大规模知识图谱预训练模型，解决 3 个重点问题：out-of-sample、inductive learning 和 logic-aware embedding。预训练模型用于亿级企业知识图谱上的链接预测、结点分类、逻辑推理等	

## 论文发表

● ( NIPS2022在投 ) FLEX: Feature-Logic Embedding Framework for CompleX Knowledge Graph Reasoning	
本文提出一种 feature-logic 的查询嵌入表示框架进行知识图谱上复杂逻辑推理。该框架是第一个不仅能处理所有一阶逻辑，还能支持多种特征空间的表示框架。该框架利用神经网络和向量逻辑建模了所有一阶逻辑，基于查询语句构建逻辑计算图，从给定的锚实体出发，沿计算图执行逻辑运算，获得查询嵌入与所有候选实体评分，最终得到答案实体。在 3 个公开数据集上性能达到 SOTA	
● ( NIPS2022在投 ) TFLEX: Temporal Feature-Logic Embedding Framework for CompleX Reasoning over Temporal Knowledge Graph	
本文提出新任务：时序知识图谱上的逻辑推理。将FLEX推广到时序知识图谱上，进一步处理Before、After、Between等时序逻辑	
● ( Findings of NAACL 2022 ) QubitE: Qubit Embedding for Knowledge Graph Completion	
本文提出量子力学启发的量子态嵌入模型进行知识图谱补全。该模型将实体嵌入为量子态，将关系嵌入为量子门，性能达到SOTA	

## 技能/证书及其他

- **技能**：数据分析可视化、算法研究 ( Python+PyTorch/自研算法库 )，移动端 ( Android 原生精通，Java/Kotlin )，前端 ( 熟悉，TypeScript/SCSS/HTML/CSS/JavaScript+React/Vue )，后端 ( 熟悉，Java/Python/Go/Rust+Spring boot全家桶 )
- **证书**：英语 ( CET-4、CET-6 )，计算机二级 ( Java )，前端开发 ( freecodecamp证书 )
- **技术开源**：Github 累计 star 2.0k，受欢迎开源项目：LaTeX\_OCR\_PRO ( 识别数学公式图像为 LaTeX, star 617 )，PythonDataMining ( 数据挖掘，star 551 )，LaTeX\_OCR ( 数学公式识别, star 296 )