



# 马学威

籍贯：山东省济宁市  
电话：(+86) 188-4454-9019  
邮箱：xuew\_ma@163.com  
博客：[maxuewei2.cnblogs.com](http://maxuewei2.cnblogs.com)  
GitHub：[github.com/maxuewei2](https://github.com/maxuewei2)

## 教育背景

2017.09 – 2020.07 吉林大学 计算机应用技术 硕士 (推免)  
2020 届优秀毕业生；连续三年获得学业奖学金及优秀奖学金  
2013.09 – 2017.07 吉林大学 计算机科学与技术 学士  
2016 年获校优秀学生；连续三年获得一等奖学金  
GPA: 3.49/4.0 GPA Ranking: 11/355

## 研究方向

图计算/图数据挖掘、网络表示学习/图嵌入、复杂网络

## 论文

<https://github.com/maxuewei2/RiWalk>

- RiWalk: Fast Structural Node Embedding via Role Identification  
Xuewei Ma, Geng Qin, Zhiyang Qiu, Mingxin Zheng, Zhe Wang  
ICDM - IEEE International Conference on Data Mining, 2019 (CCF B 类会议)  
Regular Paper, Oral Presentation, **Acceptance Rate: 9.08%**  
使用 C 和 Python 实现，比基线算法快 10 倍以上.

## 项目 (部分)

<https://github.com/maxuewei2>

- 知乎爬虫 2017.09
  - 支持多账户切换、多线程、多代理的知乎爬虫，爬取用户信息及头像、关注用户列表、关注话题、关注问题等.
  - 使用 Python 和 Java 实现，未使用任何框架. 为防屏蔽，同时爬取代理并维护代理池.
  - 用两周时间里爬取到的数据构建了一个用户数 11 万的社交网络.
- 结合知识图谱的推荐系统 2017.03 – 2017.06
  - 使用贝叶斯个性化排序优化矩阵分解，结合知识图谱提升精确度并避免冷启动.
  - 使用 C++ 实现，使用知识图谱表示学习等技术，在电影和书籍推荐任务上评测.
- 遗传算法生成正则表达式对文本分类 2017.10
  - 使用遗传算法生成正则表达式集合，使匹配较多正类文本，而尽量少匹配负类文本.
- JLUThesis2020 2020.3
  - 吉林大学硕博学位论文 LaTeX 模板.

## 技能

专业技能: (大规模) 网络表示学习/图嵌入，图数据挖掘，社交网络挖掘，机器学习，推荐系统.  
具有良好的数据结构、算法、计算机组成与体系结构、计算机网络等专业基础.

编程技能: Python, C, Java; NetworkX, PyTorch; GNU/Linux,  $\LaTeX$ .

语言: 英语 (六级 511)，可流畅阅读英文文档.

学术技能: 具有较好的英文专业文献阅读能力及写作能力.