

# 梁俊然

兴趣方向: 数据挖掘; Github: <https://github.com/kingsley0107>

电话: +852 53194591, 邮箱: [kingsley0107@gmail.com](mailto:kingsley0107@gmail.com), 个人网页: <https://kingsley0107.github.io/>

## 教育经历

### 深圳大学 (本科)

2019/09 – 2023/07

专业: 地理空间信息工程, 绩点 3.25 / 4.5, IELTS 6.5 / 9.0 (Reading 7.5, Writing 6.0)

奖项: 2021-2022 建筑与城市规划学习之星奖学金 (绩点班级排名 2/31)

毕业论文: 基于手机信令数据的高精度城市空间社会隔离量化计算

### 香港理工大学 (硕士在读)

2023/09 – 2025/07(expected)

专业: 城市信息学与智慧城市, 绩点 4.15 / 4.3

研究方向: 时空数据挖掘, 轨迹预测

毕业论文: Understanding the effects of mobile phone app usage in travellers' movement predictability

## 研究经历

### 手机 App 使用序列信息对游客轨迹预测性能影响的探究

2023/12 – Now

- **基于深度学习的信息提取:** 通过 App Encoder(LSTM based and Self-Attention based)将每一个停留点对应的 App 序列进行信息提取, 并与停留点 Embedding 输出结合得到 physical-virtual 双重维度的用户信息, 再使用 LSTM 进行物理空间的 Next-Step Location 轨迹预测
- **空间单元 Adaptive clustering:** 通过自适应的聚类算法进行空间单元划分, 解决轨迹数据在空间单元上分布极度不均匀的情况, 保证了深度学习过程中的分类样本均衡

### 基于手机信令轨迹数据的高精度城市空间社会隔离量化计算

2022/10 – 2023/05

- **信令数据轨迹处理:** 基于居住点识别算法与房价分布数据构建全样本(1%抽样)手机用户的经济相似性网络
- **任意尺度的社会隔离计算框架:** 从个体角度: 通过加权计算与其他个体'共空间'交互强度及对应的经济相似度得到个体层面的社会隔离指数; 从空间角度: 通过加权计算该空间单元中两个体的'共空间'交互强度及对应的经济相似度得到空间层面的社会隔离指数

### 基于街景图像大数据的交通事故与建成环境要素的相关性分析

2023/09 – 2023/12

- **街景图像语义分割模型耦合 XGBoost 的建模及相关性探究:** 通过图像语义分割(pre-trained)模型提取 3w+张事故街景图像的城市要素, 使用 XGBoost 与 SHAP 进行城市要素及事故发生的相关性探究, 通过消融学习分别论证城市功能(POI)数据与城市感官(街景)的边际效用

## 项目经历

### CityViewer (企业级项目)

2022/12 – 2023/03(实习)

- **基于 Scrapy 框架的数据爬取:** 使用 Scrapy 框架进行多源空间数据爬取(街景、房价、POI 等), 并使用 geopandas 等工具进行数据清洗;
- **多层次空间数据库管理:** 实现多层次的空间数据管理方式以应对空间数据结构的复杂性: 如多级 mongoDB 进行数据属性尺度及字段名称的统一, PostgreSQL 进行生产数据的管理等

### 城市路网简化算法

2023/03 – 2023/05

- **简洁的开源路网数据简化算法:** 基于缓冲区-消融-重关联的空间处理流程批量简化杂乱的網路数据(如 qq, osm 等), 同时保存路网属性信息

## 实习经历

### 北京数城未来科技有限公司, CityDNA

2022/12 – 2023/03

岗位: 数据分析实习生

主要职责: 数据爬取; 空间数据清洗、处理及入库; [CityViewer](#) 各类别数据查询的 SQL 脚本编写

技能与软件: Python, Pytorch, MongoDB, PostgreSQL, Git, Javascript, Mapbox, Echarts, Arcpy