



## 基本信息

姓名：高怡睿  
电话：18321756882  
家庭住址：广东省汕头市

民族：汉  
邮箱：gaoyirui98@gmail.com  
学历：硕士

## 教育背景

- 2019.8 – 2021.8 美国密歇根大学安娜堡分校 数据科学（硕士） GPA 4.0/4.0  
主修课程：机器学习，自然语言处理，数据分析，数据可视化，时间序列分析，贝叶斯建模
- 2016.9 – 2020.8 上海交通大学 电子与计算机（本科） GPA 3.6 / 4.0  
主修课程：数据结构，算法导论，计算机网络，计算机组成

## 项目经历

## 结合用户与子论坛特征的用于 Reddit 论坛的表情预测

- 2019.10 – 至今 密歇根大学安娜堡分校信息学院 Foreseer 实验室 项目主要承担者
- 使用 SQL 语言存储 Reddit 数据集并且做表情过滤处理；
  - 使用机器学习与深度模型，结合用户特征与子论坛特征构建特征空间，对 Reddit 数据集进行用户表情预测

## 从小说总结文本中的人物关系变化提取

- 2020.02 – 至今 密歇根大学安娜堡分校信息学院自然语言处理课程 项目承担者
- 使用 BookNLP 模型对小说总结文本中的人物关系进行提取；
  - 通过建立序列模型，对序列中句子使用句子嵌入方法，学习不同小说总结文本中人物关系变化的处理

## 对癌细胞图像的细胞分割算法研究

- 2020.02 – 至今 密歇根大学安娜堡分校信息学院机器学习课程 项目承担者
- 使用 Pytorch 构建细胞分割算法模型，如 U-Net, U-Net++ 等；
  - 通过对细胞感受野大小的调整，以及添加细胞风格分类器，来优化现有的细胞分割算法模型

## 结合情感分析的主题热度预测

- 2019.03 – 2019.08 上海交通大学电子信息与电气工程学院先进网络实验室 第二作者
- 结合情感分析手段对主题热度预测方法进行优化与合并；
  - 使用时间序列模型如 ARIMA 模型对 Twitter 数据的热度序列进行预测，从而实现热度预测结果

## 学术发表

- Li, Jinning, **Yirui Gao**, Xiaofeng Gao, Yan Shi, and Guihai Chen. "SENTI2POP: Sentiment-Aware Topic Popularity Prediction on Social Media." In *2019 IEEE International Conference on Data Mining (ICDM)*, pp. 1174-1179. IEEE, 2019.
- Wu, Qitian, **Yirui Gao**, Xiaofeng Gao, Paul Weng, and Guihai Chen. "Dual Sequential Prediction Models Linking Sequential Recommendation and Information Dissemination." In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 447-457. 2019.

## 技能

语言：英文水平较强，沟通顺畅，写作能力强  
编程语言：Python, C++/C, R, Matlab, SQL, HTML, CSS, JavaScript  
技能领域：数据科学，机器学习，数据分析

## 获得奖项

上海交通大学本科生奖学金 B 类  
上海交通大学本科生学业进步奖  
上海交通大学密西根学院学业荣誉奖 (Dean's List)

