

# 王炫宇



电话: 7343531516 | 邮箱: olivwang@umich.edu

微信: oliverwang950822 | 领英: <https://www.linkedin.com/in/xuanyu-wangx/>

个人网站: <https://github.com/oliverwangx>

求职意向: 2020年暑期实习-软件工程师

## 教育经历

密西根大学-安娜堡分校 - 数据科学 硕士 信息学院

2018年9月 - 2020年12月

相关课程: 深度学习, 机器学习, 数据库管理系统, 数据挖掘, 信息检索

上海交通大学 - 电子与计算机工程 本科 密西根学院

2014年9月 - 2019年5月

相关课程: 数据结构, 算法导论, 计算机体系架构

## 实习经历

埃森哲 - 数据分析师

2017年12月 - 2018年1月

- 为制药领域客户提供管理咨询服务, 采访3个不同核心部门客户, 提供人力资源重新配置方案
- 使用SQL和Python语法结合pandas库对数据进行空值、异常值等处理, 做平均值、箱线图等描述性分析
- 使用Python可视化工具seaborn和Matplotlib来展示每月各部门员工工作时间, 工作量等指标

密西根大学先见者实验室 - 研究助理

2019年8月 - 2020年1月

- 利用RNN深度学习网络解决问题回答问题
- 利用tweet的emoji数据来解决多义词歧异问题

## 项目经历

设计基于多进程、多线程、套接字编程的MapReduce后端框架 (Python)

2020年1月 - 2020年3月

- 利用多进程、多线程、套接字编程来模拟MapReduce框架(Master-Slave 模型), 并行处理 Map, group, reducer等任务
- 设计主节点进程(NameNode), 不同线程分别实现消息接收处理, 任务分配、作业监控、任务监控, 故障容许等功能
- 设计从节点进程(DataNode), 不同线程分别实现消息接收, 任务处理, 传输心脏信息, 节点初始化等功能
- 利用套接字来实现不同进程通信, 心跳跳动信息用UDP协议通信, 其他信息用TCP协议和JSON格式通信

迷你Instagram全栈开发项目(Python/Flask)

2019年12月 - 2020年2月

- 用Flask视图, MVC框架来开发若干社交网络功能, 包括: 认证, 授权, 关注, 取消关注, 发布动态
- 用MySQL设计关系型数据库管理后端session信息, 储存用户新的评论、动态等信息
- 用HTML, CSS, React, Jinja模版来展现前端网页
- 用RESTFUL API 来分离前后端, 用JSON传输信息

基于协同过滤的电影推荐系统(Hadoop)

2020年1月 - 2020年3月

- 拉取Netflix数据(1G 包括用户, 评分, 电影ID), 推荐用户电影
- 用物品协同过滤算法来计算评分矩阵缺失值, 推荐电影
- 用MapReduce, HDFS计算用户评分矩阵, 同现矩阵, 正则化同现矩阵, 矩阵乘法评分矩阵与同现矩阵计算电影评分
- 最终结果用RMSE计算与真实值平均差距1.32(范围为1-5分)

利用n元语法和语言模型做谷歌自动补全(Hadoop)

2019年11月 - 2019年12月

- 用MapReduce从维基文档中建立N-Gram语言库并分发数据到HDFS上
- 用MapReduce从N-Gram语言库中根据统计概率搭建语言模型并储存在MySQL数据库中
- 用HTML, CSS, JavaScript来展现前端网页功能并使用SQL调用数据实现实时自动补全