

# 王鸿儒

<https://rulegreen.github.io/>  
(+86) 181-3468-0118 ◇ [hrwang@cse.cuhk.edu.hk](mailto:hrwang@cse.cuhk.edu.hk)



## 教育背景

### 中国传媒大学

2015/09 - 2019/09

计算机科学与技术(大数据技术与应用)

GPA: 3.27/4.00

计算机操作系统, 计算机网络, 数据结构与算法, 机器学习, 自然语言处理

### 香港中文大学

2019/09 - 2020/11

计算机科学与工程

GPA: None

数据挖掘, 大数据技术与应用, 机器学习, 文本挖掘

## 实习经历

### 推荐系统算法实习生, 北京达佳互联信息科技有限公司(快手)

2019/04 - 2019/07

音乐侧:负责快手内部音乐榜单快听榜等的的数据产品, 使用grpc的proto3构建数据接口, 与服务端人员对接, 根据接口访问情况优化代码逻辑; 负责66铃声中音悦台, 伤感榜等榜单内容的数据接口; 根据运营需求, 使用jsp构建音乐流量配置中心如榜单位置, 频道位置等, 提高音乐运营的效率

推荐侧:优化音乐标签页下面相似音乐的推荐逻辑, 主要使用协同过滤与相同tag等兜底策略;优化音乐的示例视频, 使得质量较高并且发布时间较短的视频能够更好地被分发; 对原声音乐的推荐逻辑与mmu部门进行对接, 使得原声音乐与非原声音乐互相映射, 扩大推荐范围

分析侧:根据业务需求设计ab test实验, 并且使用python为ab test实验添加实验指标如日活, 月活, 观剧时长, 订阅人数等等

## 项目经验

### SemEval 2020 Task 4: CommonSense Detection and Explanation

2019/09- Now

此项目是关于检测常识与解释常识的, 具体任务包括从两个语句中选择出没有意义的句子, 以及针对没有意义的句子从已知原因中选择一个来解释, 这两个任务都是text classification问题, 第三个任务是让机器自己来解释为什么第一个任务当中选择的没有意义的句子为什么没有意义, 是一个text generation问题, 在这个项目, 使用了seq2seq, BiLSTM, self-attention等方法达到了SOTA的效果

### COMAP美国大学生数学建模比赛

2018/02- 2018/03

收集美国四大洲的能源使用数据, 建立不同的能源使用数学模型, 根据不同地区的能源分布以及地区特点, 制定符合地区发展的能源策略, 并使用神经网络预测未来20年、50年的能源储备

使用多种数学分析方法如多因子分析和控制变量分析等方法, 衡量模型准确率与健壮率, 对数据进行线性回归分析, 拟合相关数据, 使用echart.js和d3.js绘制能源配比和模型结果, 比赛获得M奖(TOP10)

### WeCampus微信小程序

2017/04 - 2018/05

组建5人团队, 前期主要通过问卷调查收集约170份在校师生对于校园活动的需求, 通过量化指标对用户需求进行聚类分析, 聚焦小程序核心功能与交互逻辑

绘制产品手稿图, 通过Xmind以及processon等工具绘制思维导图以及页面流程图, 与团队沟通页面设计, 协调前后端的进度与项目计划

编写核心功能代码如模拟登陆中验证码识别, 前后端数据交互, 使用深度学习中的CNN网络对验证码图片进行识别, 图片数据集约500张, 64\*64, 识别准确率达94%

## 荣誉奖项

互联网+创新大赛三等奖

2018/10

WeCampus软件著作权一份

2018/06

2018年美国大学生数学建模比赛一等奖

2018/02

ICBDA会议一作ICIS会议二作

2018/02