

## 黄湘云

中国矿业大学（北京）  
北京市海淀区学院路丁 11 号，100083

188 1097 2907

xiangyunfaith@outlook.com

<https://xiangyun.netlify.com/>

## 教育经历

中国矿业大学（北京）	北京市	2015.09–2018.12
理学硕士学位 理学院 统计学		
毕业论文：《空间广义线性混合效应模型及其应用》		
研究方向：数据分析与统计计算		
中国矿业大学（北京）	北京市	2011.09–2015.06
理学学士学位 理学院 数学与应用数学		

## 实习经历

北京一流科技有限公司	机器学习实习生
工作的具体内容：	
将时域和频域方法与深度学习方法应用于预测电力消费，涉及数据清洗，可视化，指数平滑，ARMA 和 GARCH 模型，贝叶斯时间序列分析工具 prophet（包含频谱分析），尖点检测，长短期记忆网络 LSTM 及 Keras 工具	
新浪公司	数据分析师实习生
工作的具体内容：	
新浪新闻客户端日志分析，主要使用 SQL 从 ClickHouse 数据仓库提取数据，R 语言分析和可视化，并完成日报，用数据分析协助其他部门决策，如服务器资源调度等。	

## 校园及社会实践

院团委宣传部干事	2011.10–2012.06
校红会学生分会组织部长	2012.06–2013.06
校红十字学生分会组织的爱心支教活动	2012.09–2012.10
大学生科研创新训练项目组长	2012.09–2014.12
CDA 数据分析师培训	2016.04–2016.06
北京大学高维统计短期课程（高维统计与网络分析）	2016.06
国家自然科学基金项目（编号：11671398）项目组成员	2017.01–2020.12
研究生助教（数学建模、数值分析、运筹学等课程）	2015.09–2016.12

## 资格证书及获奖情况

全国大学英语四级（448）和六级（462）	2011.06–2014.06
2013 年高教社杯全国大学生数学建模竞赛（北京市二等奖）	2013
2012 年本科生优秀学生奖学金（三等奖学金）	2012
2013 年本科生优秀学生奖学金（三等奖学金）	2013
第 29 届全国部分地区物理竞赛（三等奖）	2012
校科技文化节科普创作大赛（优秀奖）	2013
校物理实验竞赛（一等奖）	2013
2015-2016 年度研究生优秀学生（一等奖学金）	2016
2016-2017 年度研究生优秀学生荣誉称号	2017

## 数据可视化

### 空间数据可视化

- 案例 1: 从中国地震台网获取 2012 年 4 月至 2017 年 8 月 5 级及以上的地震数据
- 案例 2: 从美国地质调查局获取 1973 年至 2010 年世界各地 6 级及以上地震数据集
- 主要工作: 基于 ggplot2 和 animotion 等 R 包, 制作动态、平面和立体时空数据可视化

### 社区开发者网络可视化

- 分析目标: 可视化 R 语言社区中开发者相互合作的关系网络、挖掘出对社区贡献大的组织、个人
- 主要工作: 抓取 CRAN 官方 11000+ R 包信息、RStudio 官网下载日志, 正则表达式提取开发者、贡献者字段, 清洗后, 建立有向图网络, 使用 ggplot2 等可视化包, 呈现贡献关系网络。

## 项目实战

### 遗传性疾病基因数据的分析和建模

- 分析目标: 定性与定量相结合的方法找出与遗传性疾病相关的基因位点
- 使用 Pearson 卡方检验降维, 岭回归、Lasso 回归和适应性 Lasso 回归做特征提取, 利用 Cp、AICC、GCV、BIC 准则做特征选择, logistic 分布做损失函数, 得到很可能是致病位点的集合。

### 可重复的定时数据分析报告

- 在大规模新浪运维日志数据上, 基于 Docker、Clickhose 和 R Markdown 制作数据分析报告

## 专业技能

语言 R、SQL、 $\text{\LaTeX}$ 、MATLAB、Python

系统 Windows、Linux

工具 RStudio、Github、Markdown、Docker、ggplot2、SparkR

统计 数据清洗、探索、处理、建模与可视化、统计计算与绘图

英语 能够流畅地阅读专业领域内的论文和其它文献, 并熟练地用英文检索遇到的问题

## 社区贡献: 统计之都 <https://cosx.org/>

R 语言做符号计算 (一作)	<a href="https://cosx.org/2016/07/r-symbol-calculate">https://cosx.org/2016/07/r-symbol-calculate</a>
随机数生成及其在统计模拟中的应用 (一作)	<a href="https://cosx.org/2017/05/random-number-generation/">https://cosx.org/2017/05/random-number-generation/</a>
漫谈条形图 (一作)	<a href="https://cosx.org/2017/10/discussion-about-bar-graph/">https://cosx.org/2017/10/discussion-about-bar-graph/</a>
心理学的危机 (主审)	<a href="https://cosx.org/2017/09/psychology-in-crisis/">https://cosx.org/2017/09/psychology-in-crisis/</a>
在统计之都的工作	编辑、作者、审稿人、论坛版主
第 11 届中国 R 语言会议 (北京) 演讲嘉宾	2018.05.27