

大数据分析处理

Big Data Analysis and Processing

尚家兴 ○ 主讲

管梦雅 ○ 助教



主讲教师：尚家兴

姓名	尚家兴	电话	17783974208	QQ	511749656
办公室	A区主教学楼 1805	邮箱	shangjx@cqu.edu.cn		
专业技术 职务	副教授	单位	重庆大学计算机学院计算机工程系		
研究兴趣	社交网络分析与挖掘；行业大数据分析；信息传播；图神经网络				
个人履历	2018.9-至今	重庆大学 计算机学院	副教授		
	2016.2-2018. 8	重庆大学 计算机学院	讲师		
	2104.9-2015.1	香港城市大学	研究助理		
	2010.9-2016.1	清华大学自动化系	博士		
	2006.8-2010.7	清华大学自动化系	学士		
个人简介	http://www.cs.cqu.edu.cn/info/1320/4108.htm				

姓名：管梦雅

电话：150-0867-8592

QQ：2091344188

邮箱：2091344188@qq.com

职责：协助批阅作业、辅导实验，帮助同学们答疑解惑。

扫码加入

请将个人名片修改为
“班级-姓名” 格式

例如：

计科02-张三

物联网01-李四



大数据分析处理-2...

群号：849811521



扫一扫二维码，加入群聊。

基本信息

- **课程性质**：专业选修课
- **学分**：**2.5**学分（**48**学时=**32**课堂学时+**16**实验学时）
- **学习方式**：课堂学习/讨论/报告（**16**次）+实验（**4**次）
- **基础课程**：计算机导论、Python程序设计基础、概率论、数据库

课程目标

- 系统、完整地掌握大数据分析处理的基本概念与方法，具有深入学习和研究大数据相关技术的能力。
- 利用Anaconda, PyTorch等开源工具平台进行实验，加深知识点理解，促进融会贯通、学以致用。
- 针对复杂工程问题，设计并完成课程项目，并通过PPT进行项目展示，提升解决实际大数据应用问题的能力以及知识表达交流能力。
- 具备大数据文献查找、分析、综述和应用的能力。

教材

[1] 《大数据分析》，程学旗 主编，高等教育出版社，2019年4月出版

参考书

1. 《数据挖掘：概念与技术(原书第3版)》
[美] Jiawei Han, [美] Micheling Kamber,
[美] Jian Pei 等著，范明，孟小峰 译，机械工业出版社，2012年8月出版.
2. 《统计学习方法（第2版）》，李航 著，清华大学出版社，2019年5月出版
3. 《大数据处理》，金海，石宣化 著，高等教育出版社，2018年12月出版
4. 《大数据导论》，梅宏 著，高等教育出版社，2018年11月出版



考核方式

- 平时作业：20%
- 实验（4次）：20%
- 课程答辩（课堂答辩PPT制作及讲解）：20%
- 课程报告：40%

课堂纪律

- 课程作业、实验报告和期末报告须按时提交，迟交者将酌情扣分。
- 课程作业、实验报告和期末报告严禁抄袭，一旦发现，抄袭者和被抄袭者本次成绩均按零分计算
- 旷课：旷课一次总分扣10分，特殊情况提前向老师请假。

大数据导论

- 侧重大数据的体系结构与框架介绍
- 大数据整体框架：Hadoop、Spark
- 大数据存储框架：HDFS、HBase
- 大数据计算框架：MapReduce、YARN
- 流式计算框架：Spark Streaming、Storm

会使用框架！

VS

会设计算法！

大数据分析处理

- 侧重大数据分析与处理的方法和理论介绍
- 数据的认识与理解
- 大数据机器学习方法
- 大数据算法理论
- 文本、知识、社交媒体、网络大数据的分析理论与方法

Any question?