

2019 版数据科学与大数据技术本科专业培养方案

一、专业简介

数据科学与大数据技术专业以“计算机科学与技术”和“软件工程”两大学科为基础，依托“医疗大数据”实验室和各大附属医院的数据资源，专业涵盖大数据采集、存储、处理、分析以及应用等相关研究。注重培养学生的动手能力、分析能力和国际视野。本专业师资队伍数量充足，专业、职称、学历、年龄结构合理，已建立起一支年青的、科研能力较强且教学水平较高的优秀教师队伍。

二、培养目标

本专业培养适应我国社会经济发展需要，德、智、体、美、劳全面发展，具备扎实的人文科学、自然科学和工程技术的基础知识，掌握大数据科学与技术所需要的计算机、网络、数据处理等相关学科的基本理论、专业知识、基本技能及专业发展动态，具备大数据采集、存储、处理与分析、传输与应用等专业技能，能在科研院所、企业、高新技术公司从事大数据系统的研究、开发、部署与应用以及大数据处理与分析等工作，培养重德行、乐奉献、厚基础、强能力、高素质，具有创新精神、国际视野和较强实践能力的高级专门人才。

预期五年以上的毕业生：

目标 1：能在知名企业、高新技术公司担任数据科学家、数据分析师等职位；

目标 2：能在政府单位、科研院所、企业担任大数据平台架构工程师等职位；

目标 3：能获得研究生学位，取得大数据相关的科研成果；

目标 4：能在企业和单位担任大数据平台开发工程师和运维工程师等。

三、培养要求

本专业培养掌握大数据科学与技术相关的基本理论和基本知识，系统地掌握数据科学与工程专业知识，具备大数据应用系统设计与开发、大数据处理与分析能力，以及一定的科研工作能力，达到知识、能力与素质的协调发展。

（一）本专业培养的人才应具备如下知识、能力和素质要求：

1. 知识要求

① 工具性知识：具有外语、文献检索、科技写作的能力。

② 人文社会科学知识：具有基本的文学、哲学、政治学、社会学、法学、心理学、思想道德、

职业道德、艺术、军事和经济管理等方面的素养。

③ 基础知识：具有数学、物理学方面的基础知识。

④ 专业基础知识：具有计算机科学与技术的基础知识，涉及高等数学、程序设计、数据结构、数据库、统计学、算法分析与设计等。

⑤ 专业知识：具有数据存储、集成与处理，大数据分析与管理，大数据平台构建与维护等相关的知识。

2. 能力要求

① 获取知识能力：掌握文献检索、资料查询及运用现代信息技术获取相关信息的基本方法；具有终身学习与适应发展的能力。

② 应用知识能力：具有将数据科学与大数据技术的基础知识、基本方法和工具应用于大数据存储、集成与处理，大数据平台构建、管理维护的基本能力；具有初步的医学大数据项目管理等基本能力，具有良好的工程素养，能够在大数据相关工程实施过程中选择和使用合适的工具，在系统设计、开发、集成等活动中能充分考虑对社会、经济、环境等的影响。

③ 组织合作能力：具有一定的组织管理能力、较强的表达能力和人际交往能力；具有国际视野和跨文化的交流、竞争与合作能力。

④ 创新能力：在科学研究、工程设计和实践等方面具有一定的创新意识和能力。

3. 素质要求

① 思想道德素质：热爱祖国，拥护中国共产党的领导，树立科学的世界观、人生观和价值观，具有责任心和社会责任感；具有法律意识，自觉遵纪守法；热爱本专业，注重职业道德修养；具有诚信意识和团队精神。

② 文化素质：具有一定的文学艺术修养、人际交流与沟通能力和现代意识。

③ 专业素质：掌握科学思维方法和工程设计方法，具备良好的工程素养；具有创新意识和创业精神；具有严谨的科学态度和务实的工作作风。

④ 身心素质：具有较好的身体素质和心理素质。

(二)本专业对学生的毕业要求具体内容如下：

1. 工程知识：能够将数学、自然科学、工程基础和专业知用于解决复杂数据科学与大数据技术问题。

2. 问题分析：能够应用数学、计算机工程科学、医学的基本原理，识别、表达、并通过文献研究分析复杂数据科学与大数据技术问题，以获得有效结论。

3. 设计/开发解决方案：能够针对大数据领域特定需求设计搭建复杂大数据平台与分析系统，并能够在设计环节中体现创新意识，考虑社会、健康、安全、法律、文化以及环境等因素。

4. 研究：能够基于科学原理并采用科学方法对大数据系统的设计与开发进行研究，包括需求分析、数据获取与说明、系统前后端设计等、并通过信息综合得到合理有效的结论。

5. 使用现代工具：能够针对大数据分析与管理等复杂工程问题，选择与使用恰当的框架平台、数据存储与处理技术、管理工具和其它信息技术工具，包括对复杂工程问题的预测与模拟，并能够理解其局限性。

6. 工程与社会：能够基于工程相关背景知识进行合理分析，评价数据科学与大数据技术实践和复杂工程问题解决方案对社会、健康、安全、法律以及文化的影响，并理解应承担的责任。

7. 环境和可持续发展：能够理解和评价针对复杂数据科学与大数据技术问题的专业工程实践对环境、社会可持续发展的影响。

8. 职业规范：具有人文社会科学素养、社会责任感，能够在工程实践中理解并遵守工程职业道德和规范，履行责任。

9. 个人和团队：能够在多学科背景下的团队中承担个体、团队成员以及负责人的角色。

10. 沟通：能够就复杂数据科学与大数据技术问题与业界同行及社会公众进行有效沟通和交流，包括撰写报告和设计文稿、陈述发言、清晰表达或回应指令。并具备一定的国际视野，能够在跨文化背景下进行沟通和交流。

11. 项目管理：理解并掌握数据科学与大数据技术管理原理与经济决策方法，并能在多学科环境中应用。

12. 终身学习：具有自主学习和终身学习的意识，有不断学习和适应发展的能力。

四、主干学科

计算机科学与技术；软件工程

五、核心知识领域与专业核心课程

核心知识领域：数学与统计（MS）、程序设计基础（PF）、计算机系统（CS）、数据库系统（DB）、算法设计与分析（AA）、数据处理（DP）、数据分析（DA）、云计算技术（CC）、人工智能（AI）。

专业核心课程：计算思维、离散数学、数据结构、操作系统、数据库原理、计算机网络原理、算法设计与分析、大数据平台及编程实践、大数据分析与管理、Python数据处理、数据可视化、分布式系统、机器学习。

六、主要实践性教学环节与主要专业实验

实验、实训、课程设计、认识实习、生产实习、毕业实习、毕业设计
七、学制与学位

学 制：标准学制为 4 年，学习年限为 3-8 年

学分要求：学业学分 172 学分，第二课堂 10 学分

授予学位：工学学士

八、各类课程学分学时分配表

课程性质	课程属性	学分	占学分比例	教学学时	占课程学时比例
公共基础课平台	必修	29	16.86	528	21.78
公共基础课平台	选修	8	4.65	128	5.28
学科基础课平台	必修	52.5	30.52	920	37.95
学科基础课平台	选修	25	14.53	400	16.5
专业课平台	必修	11.5	6.69	192	7.92
专业课平台	选修	16	9.3	256	10.56
集中性实践环节	必修	30	17.44		
总计		172	100	2424	100

九、各平台课程设置与学分

(一) 公共基础必修课平台

课程编号	课程名称	考核	学分	总学时	讲课	实验	实践	学期	责任单位	备注
305000 00510	大学生职业发展与就业指导1 Career Development and Employment Guidance for College students	考查	0.5	6	2	4	0	1	创新创业基础教研室	
518000 00210	大学体育1 University physical education1	考查	1	32	4	28	0	1	基础体育教研室	
516000 28110	大学英语A1 College English A1	考试	3	48	48	0	0	1	大学英语第一教研室	

517000 00510	形势与政策1 Situation and Policy	考查	0.5	8	6	2	0	1	形势与政策教研室
517000 00410	中国近现代史 纲要 The Conspectus Of Chinese Modern History	考查	3	48	40	8	0	1	中国近现代史纲要教研室
518000 00310	大学体育2 University physical education2	考试	1	32	4	28	0	2	专项体育教研室
516000 28210	大学英语B2 College English B2	考试	3	48	48	0	0	2	大学英语第一教研室
517000 00310	思想道德修养 与法律基础 Education of Ideology and Morality and Introduction to the Law	考试	3	48	40	8	0	2	思想道德修养与法律基础教研室
517000 00610	形势与政策2 Situation and Policy	考查	0.5	8	6	2	0	2	形势与政策教研室
518000 00410	大学体育3 University physical education3	考查	1	32	4	28	0	3	基础体育教研室
516000 63210	大学英语3 College English 3	考查	2	32	32	0	0	3	大学英语第一教研室
517000 00110	马克思主义基本 原理概论 An Outline of fundamental Principles of Marxism	考试	3.5	56	48	8	0	3	马克思主义基本原理教研室
517000 00710	形势与政策3 Situation and Policy	考查	0.5	8	6	2	0	3	形势与政策教研室
518000 00510	大学体育4 University physical education4	考试	1	32	4	28	0	4	基础体育教研室
517000 00210	毛泽东思想和 中国特色社会 主义理论体系	考试	4.5	72	64	8	0	4	毛泽东思想与中国特色社会主义理论体系概论教研

	概论 Introduction of Mao Zedong Thought and Socialism Theory System with Chinese Characteristi cs								室	
517000 00810	形势与政策4 Situation and Policy	考查	0.5	8	6	2	0	4	形势与政策教研 室	
305000 00710	大学生职业发 展与就业指导2 Career Development and Employment Guidance for College students	考查	0.5	10	2	8	0	6	创新创业基础教 研室	
小计			29	528	364	164	0	学分要求:		29

(二) 公共基础选修课平台

课程编号	课程名称	考核	学分	总学时	讲课	实验	实践	学期	责任单位	备注
小计								学分要求:		8

公共基础选修课共计 8 学分，分 2 个模块。

限选：

1. 大学生心理健康教育与指导（1 学分，学生须在第一学期修读）；
2. 公共艺术类选修课程（2 学分，在影视鉴赏、音乐鉴赏、舞蹈鉴赏、书法鉴赏、戏剧鉴赏、戏曲鉴赏、美术鉴赏、古代名剧鉴赏等 8 门课程中选修 2 门）。

任选：

每个学生要求跨学科修读其他学科门类通识课程 5 学分；其中工科类学生要求选修至少 1 学分经济管理类通识课。

8 个学分在第五学期以前修完，第一、第三学期各修 1 学分，第二、第四、第五学期各修 2 学分。

(三) 学科基础课平台必修课

课程编号	课程名称	考核	学分	总学时	讲课	实验	实践	学期	责任单位	备注
------	------	----	----	-----	----	----	----	----	------	----

505000 14920	C++程序设计 C++ Programming	考试	5	80	80	0	0	1	医学信息工程 系	
505000 15020	C++程序设计实 验 C++ Programming Experiment	考查	1	32	0	32	0	1	医学信息工程 系	
510000 11420	高等数学A1 Advanced Mathematics A1	考试	5	80	80	0	0	1	高等数学教研 室	
505000 03420	离散数学 Discrete Mathematics	考试	4	64	64	0	0	1	软件工程系	
505000 10020	Java程序设计 实验 Java Programming Experiment	考查	1	32	0	32	0	2	网络工程系	
510000 07810	大学物理B College Physics B	考试	3.5	56	56	0	0	2	物理教研室	
510000 11520	高等数学A2 Advanced Mathematics A2	考试	5.5	88	88	0	0	2	高等数学教研 室	
505000 13620	数字逻辑 Digital Logic	考查	3	48	38	10	0	2	物联网工程系	
510000 07910	物理实验B Physical Experiment B	考查	1	32	0	32	0	2	物理实验室	
510000 08610	线性代数A linear algebra A	考试	2.5	40	40	0	0	2	工程数学教研 室	
510000 00820	概率论与数理 统计A Probability theory and mathematical statistics A	考试	4	64	64	0	0	3	工程数学教研 室	
505000 13020	计算机组成原 理 Computer Organization	考试	4	64	64	0	0	3	物联网工程系	
505000	计算机组成原	考查	1	32	0	32	0	3	物联网工程系	

13120	理实验 Computer Organization Experiment									
505000 01830	数据结构 Data Structure	考试	3.5	56	56	0	0	3	计算机科学教研室	
505000 01930	数据结构实验 Data Structure Experiments	考查	0.5	16	0	16	0	3	计算机科学教研室	
505000 23430	数据科学导论 Introduction to Data Science	考查	1	16	16	0	0	3	医学信息工程 系	
505000 00730	操作系统原理 Principle of Operation System	考试	3	48	48	0	0	4	计算机科学教研 室	
505000 10320	计算机网络原 理 Principle of Computer Network	考试	3.5	56	56	0	0	4	网络工程系	
505000 10420	计算机网络原 理实验 Principle of Computer Network Experiment	考查	0.5	16	0	16	0	4	网络工程系	
小计			52.5	920	750	170	0	学分要求:		52.5

(四) 学科基础课平台选修课

课程编号	课程名称	考核	学分	总学时	讲课	实验	实践	学期	责任单位	备注
505000 03320	计算机科学导 论 Introduction to Computer Science	考查	2.5	40	30	10	0	1	软件工程系	限选
505000 17330	计算机类专业 导论 Introduction to computer- related majors	考查	1	16	16	0	0	2	数字媒体技术 系	限选
505000 00530	Python语言程 序设计	考查	3	48	32	16	0	3	计算机科学教研 室	限选

	Python Programming									
505000 08630	Web前端技术 Web Front-end Technology	考查	2	32	20	12	0	4	数字媒体技术系	限选
505000 01030	编译原理 Compilation Principle	考试	3	48	36	12	0	4	计算机科学教研室	任选
505000 16930	分布式计算与并行计算 Distributed computing and parallel computing	考查	3	48	32	16	0	4	医学信息工程系	限选
505000 04630	数据库原理A Database Principles A	考试	3	48	40	8	0	4	软件工程系	限选
505000 04830	算法分析与设计 Algorithm Analysis and Design	考查	2.5	40	30	10	0	4	软件工程系	限选
513000 02220	卫生统计学B Health Statistics B	考查	3	48	32	16	0	4	社会医学与卫生事业管理教研室	任选
502000 04820	信号与系统B Signals and Systems B	考查	2	32	32	0	0	4	电气工程及其自动化系	限选
505000 12730	传感器技术及应用 Sensor Technology and Application	考查	3	48	32	16	0	5	物联网工程系	限选
505000 14820	现代通信原理 Modern Communication Principles	考试	3	48	48	0	0	5	物联网工程系	限选
小计			31	496	380	116	0	学分要求:		27

(五) 专业课平台必修课

课程编号	课程名称	考核	学分	总学时	讲课	实验	实践	学期	责任单位	备注
505000 09920	Java程序设计 Java Programming	考查	3	48	48	0	0	2	网络工程系	

505000 15830	机器学习 Machine learning	考查	3	48	48	0	0	4	医学信息工程系	
505000 15530	大数据平台及编程实践 Big Data Platform and Programming Practice	考查	3	48	32	16	0	5	医学信息工程系	
505000 23330	数据可视化 Data Visualization	考查	2	32	32	0	0	6	医学信息工程系	
505000 23530	数据可视化实验 Data Visualization Experiment	考查	0.5	16	0	16	0	6	医学信息工程系	
小计			11.5	192	160	32	0	学分要求:		11.5

(六) 专业课平台选修课

课程编号	课程名称	考核	学分	总学时	讲课	实验	实践	学期	责任单位	备注
505000 15630	分布式系统 Distributed system	考查	2	32	32	0	0	4	医学信息工程系	任选
505000 03830	软件工程概论 Introduction to Software Engineering	考查	2	32	24	8	0	4	软件工程系	限选
505000 01330	机器视觉与图像处理 Machine Vision and Image Processing	考试	3	48	48	0	0	5	计算机科学教研室	限选
505000 22030	深度学习 Deep Learning	考查	2	32	32	0	0	5	医学信息工程系	任选
505000 05730	数字图像处理A Digital Image Processing A	考试	3	48	32	16	0	5	数字媒体技术系	任选
505000 22930	虚拟现实技术 Virtual	考查	2	32	24	8	0	5	医学信息工程系	限选

	Reality Technology									
513000 07420	医学统计学 Medical Statistics	考查	3	48	32	16	0	5	社会医学与卫生事业管理教研室	限选
505000 15130	Python数据处理 Data processing with Python	考查	2	32	32	0	0	6	医学信息工程系	限选
505000 21730	R语言 R Programming Language	考查	2	32	24	8	0	6	医学信息工程系	限选
505000 24730	大数据与深度学习 Big data and Deep Learning	考查	2	32	32	0	0	6	医学信息工程系	限选
505000 23030	数据隐私与安全 Data privacy and security	考查	1	16	16	0	0	6	医学信息工程系	任选
505000 21830	医学大数据处理 Medical BigData Processing	考试	3	48	34	14	0	6	医学信息工程系	任选
505000 22830	智能搜索引擎技术 Intelligent Search Engine Technology	考查	3	48	28	20	0	6	医学信息工程系	任选
小计			30	480	390	90	0	学分要求:		18

(七) 集中性实践教学环节

课程编号	课程名称	考核	学分	总学时	讲课	实验	实践	学期	责任单位	备注
105000 00110	军事技能 Military Skills	考查	2	2	0	0	2	1	国防军事教研室	
505000 05330	计算机类认识实习 Computer Awareness	考查	1	1	0	0	1	2	数字媒体技术系	

	Practice									
505000 18830	Java软件开发 实训 Java Software Development Practical Traning	考查	1	1	0	0	1	3	软件工程系	
505000 00630	Python语言程 序设计实训 Python Programming Trill	考查	1	1	0	0	1	4	计算机科学教研 室	
505000 08730	Web前端技术 课程设计 Curriculum Design of Web Front- end Technology	考查	1	1	0	0	1	4	数字媒体技术 系	
505000 00830	操作系统原理 课程设计 Principle of Operation Systems Course Exercise	考查	1	1	0	0	1	4	计算机科学教研 室	
505000 15730	分布式系统课 程设计 Course Design for Distributed system	考查	1	1	0	0	1	4	医学信息工程 系	
505000 15930	机器学习课程 设计 Course Design for Machine learning	考查	1	1	0	0	1	4	医学信息工程 系	
505000 04730	数据库原理A 课程设计 Database Principles A Curriculum Design	考查	1	1	0	0	1	4	软件工程系	
305000 00810	创新创业实践 (创新创业基 础) Innovation and	考查	2	2	0	0	2	5	创新创业基础教 研室	

	entrepreneurship practice (Foundation for Innovation and entrepreneurship)									
505000 24230	大数据与深度学习课程设计 Design course of Big data and Deep Learning	考查	1	1	0	0	1	5	医学信息工程系	
505000 77230	机器视觉和图像处理课程设计 Course Design of Computer Vision and Image Processing	考查	1	1	0	0	1	5	计算机科学教研室	
505000 18530	计算机学科竞赛 Computer Science Challenge Cup	考查	2	2	0	0	2	5	软件工程系	
505000 22130	深度学习课程设计 深度学习课程设计	考查	1	1	0	0	1	5	医学信息工程系	
505000 15230	Python数据处理课程设计 Course Design for Data processing with Python	考查	1	1	0	0	1	6	医学信息工程系	
505000 23730	大数据毕业实习 Graduation Practice for Data Science and BigData Technology	考查	1	1	0	0	1	6	医学信息工程系	
505000	大数据分析	考查	1	1	0	0	1	6	医学信息工程	

15430	挖掘课程设计 Course Design for Big Data Analysis and Mining								系	
505000 23930	大数据生产实 习 Production Practice for Data Science and BigData Technology	考查	2	2	0	0	2	6	医学信息工程 系	
505000 21430	大数据毕业设 计1 Graduation Design for Data Science and BigData Technology 1	考查	4	4	0	0	4	7	医学信息工程 系	
108000 00110	劳动教育 Labor education class	考查	2	2	0	0	2	7	学生工作部教务 办	
505000 23630	大数据毕业设 计2 Graduation Design for Data Science and BigData Technology 2	考查	10	10	0	0	10	8	医学信息工程 系	
小计			38	38	0	0	38	学分要求:		38

十、学期开课门数统计表

学期	课程总门数	必修课门数	选修课门数	必修课学分	选修课学分
第一学期	11	10	1	25	2.5
第二学期	13	12	1	28	1
第三学期	11	10	1	21	3
第四学期	17	9	8	18	16.5
第五学期	9	1	8	3	18
第六学期	8	5	3	6	6

概率论与数理统计A		M	L	M								
操作系统原理	M				M	M						
计算机网络原理	M				M	M						
数据结构	M				H	M						
C++程序设计	M				H	H						
数据库原理A	M				M							
数据科学导论	M			M	M							
生物信息学导论		M										
现代通信原理	M				M							
数字图像处理A	M				H	M						
人工智能导论	M				H	M						
机器学习		M		H	H							
大数据平台及编程实践		M		H	H							
分布式系统		M		H	H							
数据可视化		M		H	H							
信号与系统	M				M	M						
算法分析与设计	M				M	M						
Python数据处理		M		H	H							
分布式计算与并行计算		M		H	H							
云计算技术		M		H	H							
云存储		M		H	H							
数据分析与挖掘		M		H	H							
数据科学与大数据技术生产实习								M	M	M		
数据科学与大数据技术毕业设计								M	M	M		
操作系统原理课程设计	M				M	M						
数据库原理A课程设计	M				M	M						
C++程序设计实验		M	M		M							
C++程序设计实训		M	M		M							

十二、推荐大学科门类中英文经典必读书目

序号	书名	主编	出版社	时间（年）
1	大数据时代（第一版）	盛杨燕，周涛 译	浙江人民出版社	2013
2	深度学习	Ian Goodfellow等著	人民邮电出版社	2017
3	HBase实战（第一版）	谢磊 译	人民邮电出版社	2013
4	Java并发编程实战	童云兰 等 译	机械工业出版社	2012
5	这就是搜索引擎：核心技术详解（第一版）	张俊林 著	电子工业出版社	2012
6	GAN：实战生成对抗网络（第一版）	Kuntal Ganguly 著	电子工业出版社	2018
7	从Paxos到Zookeeper分布式一致性原理与实践	倪超 著	电子工业出版社	2015
8	数据算法：Hadoop/Spark大数据处理技巧	苏金国，杨健康 译	中国电力出版社	2016
9	深度学习入门：基于Python的理论与实现	斋藤康毅 著，陆宇杰 译	人民邮电出版社	2018
10	解析Python网络爬虫：核心技术、Scrapy框架、分布式爬虫	黑马程序员 著	中国铁道出版社	2018
11	强化学习（第一版）	赵地 译	机械工业出版社	2018
12	ZooKeeper：分布式过程协同技术详解	Flavio Junqueira 著	机械工业出版社	2016
13	Java核心技系列：Java虚拟机规范（Java SE 8版）（第8版）	爱飞翔，周志明 等 译	机械工业出版社	2016
14	Spark快速大数据分析	王道远 译	人民邮电出版社	2015
15	机器学习实战：基于Scikit-Learn和TensorFlow	奥雷利安·杰龙（Aurélien Géron） 著	机械工业出版社	2018
16	Python深度学习（第一版）	弗朗索瓦·肖莱（Francois Chollet） 著	人民邮电出版社	2018
17	Mahout实战（第一版）	王斌，韩冀中，万吉 译	人民邮电出版社	2014
18	精通Python爬虫框架Scrapy0	李斌 译	人民邮电出版社	2018

19	《高等数学（第六版）》	同济大学数学系	高等教育出版社	2007
20	概率论与数理统计（第四版）	盛骤、谢式千、潘承毅	高等教育出版社	2008

制定人：罗凌云 负责人：罗凌云 审核人：刘振宇