

专业简介

专业英文名: Data Science and Big Data Technology

专业代码: 080910T

学科门类: 工学

设置年份: 2019 年

依托学科: 计算机科学与技术

优势专业类型: 国家特色专业 国家综合改革试点专业 北京市特色专业
 卓越农林人才培养计划改革试点专业

专业认证: 是 否

专业简介:

数据科学与大数据技术专业旨在培养能够从事林业、工业等领域大数据相关的工作, 熟练掌握大数据相关技术与核心技能, 具有较强实践和创新能力的高等工程应用型人才。通过培养将使学生掌握数据科学与大数据技术所需要的计算机、网络、统计分析、大数据处理、大数据技术等相关的基本理论和知识, 熟练掌握大数据采集、集成、存储、处理、分析与应用等技术, 具备大数据工程项目的系统集成能力、应用软件设计和开发能力, 具有大数据工程师的基本能力与素质及大数据科学研究能力。

学生毕业后可适应各行业的信息化及数据管理需求, 能在高新技术企业、公司、事业单位、科研单位、国家政府机关、银行、通信、金融等部门中从事大数据采集、管理、分析、挖掘等方面工作。

数据科学与大数据技术专业学制四年, 学生毕业后授予工学学士学位。

数据科学与大数据技术专业本科培养方案

一、培养目标

本专业面向国家行业信息化发展对林业、工业大数据研究与应用的需求，坚持“理论学习+工程训练+新技术应用”的人才培养模式，培养能够从事林业、工业等领域大数据相关的工作，熟练掌握大数据相关技术与核心技能，具有较强实践和创新能力的高等工程应用型人才。专业培养将使学生掌握数据科学与大数据技术所需要的计算机、网络、统计分析、大数据处理等相关学科的基本理论和基本知识，熟练掌握大数据采集、存储、处理、分析与应用等技术；具备大数据工程项目的系统集成能力、应用软件设计和开发能力，具有一定的大数据工程师的基本能力与素质，及大数据科学研究能力。

本专业以学生实际动手能力为主导，项目式教学为驱动，理论、技术与实践课程一体化教学为主，将使学生在知识体系方面打下扎实的有大数据应用开发、大数据分析挖掘、大数据系统运维等技术的基础；在应用能力方面运用面向对象程序设计与开发技术，进行基于大数据平台软件开发及算法实现，并能够熟练的运用数据可视化工具进行数据分析结果、展示的能力。本专业通过提供丰富的课程体系和实践项目，使学生具备扎实的数据统计分析与大数据科学技术能力，掌握数据挖掘与数据分析的基本理论、基本技能及综合应用数据平台的分析技术，能轻松胜任基于企业级的大数据分析与应用项目。

本专业将围绕国家生态文明建设战略，以林业大数据分析处理的重大需求为导向，既重视学习面向大数据应用的统计学以及计算机科学与技术等学科基础知识，又强调训练面向大数据挖掘与分析技能。本专业将构建大众教育和精英教育相结合、工程型和研究型培养相结合、专业技能和领域知识相结合的教学培养体系，使学生掌握计算机科学的基本理论和方法，具备基本的实践技能，具有良好的科学素养和行业应用技能。

本专业在计算机科学与技术专业的课程体系架构的基础上，强调配置大数据分析处理相关的核心课程，并将专业选修模块分为大数据应用类与系统开发类两个子模块。

二、培养方式

本专业人才培养方式主要包括课堂理论和实验教学、实习实践教学、课程设计、毕业设计、大学生科技创新、大学生素质拓展计划等。

在课程教学中充分发挥教师主导、学生主体的作用，注意发挥学生的主动性、自觉性和创新性，更多地采用启发式、混合式、研讨式的教学方式，培养学生自主学习的能力。注重学生的设计实践教学环节，加强对学生创新能力的培养。在培养环节中逐步推行实践技能型人才培养，加强学生的自主获取新技术能力、图形图像程序设计、虚拟交互技术和数字媒体系统开发能力的训练和培养。依托大学生科技创新，推行导师制，强调个性化培养，力求为学生提供最适合的学习、发展的空间。

鼓励高年级学生在专业教师指导下参加高水平的学科专业竞赛、参与教师的科研项目，在实践中发现、弥补理论学习中的不足，开阔专业视野、培养团队精神、创新精神、科研能力和综合素质。

三、依托学科和专业核心课程

1. 依托学科：计算机科学与技术。

2. 专业核心课程：专业概论、程序设计基础、计算机导论、离散数学、电路与电子学基础、概率论与数理统计 B、数字电子技术 A、计算机组成原理 A、数据结构 A、Java 程序设计、操作系统 A、计算机体系结构、数据库系统、Linux 应用、大数据技术基础、计算机网络 A、探索性数学分析、NoSQL 数据库原理、大数据采集与集成、Python 应用、分布式系统与云计算、机器学习、Hadoop 集群计算原理及编程、时间序列分析 B、实时大数据处理技术、林业大数据分析决策、自然语言处理、数据可视化、大数据专业实践。

四、主要实践教学环节

实践教学环节主要由课程实习、综合实习、科研训练、毕业论文（设计）等组成，具体包括数据结构 A 课程设计、操作系统 A 课程设计、大数据专业实践、毕业论文（设计）。

大数据专业实践要求学生进入老师或公司的项目组，熟悉实际项目的业务流程和软件开发过程；或者按照指导老师的要求实现一个中小型的系统。通过专业实践，提高学生的动手能力，综合运用所学的专业知识，提高分析问题和解决问题的能力。

毕业论文（设计）实行导师制，具体要求由导师制订。通过毕业论文，使学生运用所学的知识、理论解决实际问题，提高思考问题、分析问题的能力，能够培养学生独立从事科学研究的能力和工程设计的能力，培养学生通过文献检索、资料查询获取信息的能力。

专业实验包括专业必修课实验和专业选修课实验。

其中，专业必修课主要课程的相关实验包括：程序设计基础实验、计算机组成原理 A 实验、数据结构 A 实验、Java 程序设计实验、大数据技术基础实验、操作系统 A 实验、计算机体系结构实验、计算机网络 A 实验、数据库系统实验、分布式系统与云计算实验、探索性数据分析、NoSQL 数据库原理实验、Linux 应用实验、机器学习实验、Hadoop 集群计算原理与编程实验、实时大数据处理技术实验等。

专业选修课共开设分 2 个模块，即大数据应用类模块、系统开发类模块。

五、毕业生应具有的知识、能力、素质

1. 具有良好的职业道德和较高的职业素养、有社会责任感；
2. 具备大数据工程师的基本素质和能力，能够从事大数据工程系统的建设、管理和维护等工作；
3. 具备基本的数据分析能力，能够在林业、信息产业以及其他各行业从事大数据处理、大数据分析等工作；
4. 具备应用大数据的基本能力，掌握数据采集、存储、处理和分析的知识与技能，具有熟练应用计算机处理和分析数据的工程能力；
5. 具备创新意识，掌握基本的创新方法，具备大数据科学研究能力。

六、学制

学制四年。

七、毕业与学位

达到本专业培养目标及相关要求，修满本专业规定学分，毕业论文（设计）合格，准予毕业。该专业毕业生至少修满 174 学分，其中必修课内讲课、必修课内研讨和专业选修共 116 学分，必修实践环节 48 学分。

达到授予学位条件的，授予工学学士学位。

八、专业教学计划表

数据科学与大数据技术专业教学计划表

课程类别	课程代码	课程名称	课内学时总计	课内学时			实习实践(周)	总学分	各学期学时分配								承担单位
				讲课	研讨	实验			一	二	三	四	五	六	七	八	
公共选修课	公共选修课分为面授课和视频课,最低选修7学分,具体要求如下: (1)面授课:每门1学分,最低选修4学分,分为人文科学、社会科学、数学与自然科学、艺术审美四大类,每类至少选修1门; (2)视频课:最低选修3学分,分为两类:第一类是精品在线开放课程,要求至少选修1门,课程名单及学分认定标准以相应学期公布的选课通知为准,如该类课程累计选修学分大于或等于3学分,可免修第二类视频课;第二类是学校认证的视频课,每门认定1学分。																
	18001650	中国近现代史纲要	44	36	8	0	0.25	3		44 (0.25)							
	18001660	思想道德修养与法律基础	40	32	8	0	0.5	3	40 (0.5)								马院
	20004320	习近平新时代中国特色社会主义思想概论	32	28	0	4		2	32								马院
	18001630	马克思主义基本原理概论	44	36	8	0	0.25	3			44 (0.25)						马院
	18001640	毛泽东思想和中国特色社会主义理论体系概论	64	52	12	0	1	5		64 (1.0)							马院
	18001671-6	形势与政策	48	24	12	12		3	8 [2]	8 [2]	8 [2]	8 [2]	8 [2]	8 [2]			马院
	18000211-2	大学英语	96	96	0	0		6	48	48							外语院
	17001041	体育	144	104	0	40		4	36	36	36	36					体育部
	17001042-4Z								[10]	[10]	[10]	[10]					
	15005750Z	管理学基础	32	28	4	0		1			32						经管院
	18000220Z	林学概论 A	32	32	0	0		1		32							林学院
	18000250	创业基础	32	32	0	0		2					32				教务处
	20000390	大学生心理健康	8	6	2	0		0.5	8								人文院
	英语模块课	分为中国文化、西方文化、学术英语、实用英语四个模块,每个模块选修1门课,共计4学分。	96	96	0	0		4			24	24	24	24			外语院
暑期学期	大一至大三暑假,学生须至少选修3学分暑期学期课程。																
学科基础教育平台(必修)	15021470	专业概论	8	8	0	0		0.5	8								信息院
	19001670	程序设计基础	48	40	0	8		3	48								信息院
	19001680	程序设计基础(课程设计)					1	1	(1)								信息院
	19001690	计算机导论	24	24	0	0		1.5	24								信息院
	15004541-2	高等数学 A	176	160	16	0		11	88	88							理学院
	15017880	线性代数 A	48	42	6	0		3	48								理学院
	18001850	物理学 D	64	64	0	0		4	64								理学院
	15017530	物理学实验 C	32	0	0	32		2	32								理学院

北京林业大学 2020 级本科人才培养方案-信息学院

	15003050	电路与电子学基础	48	42	0	6		3		48						工学院	
	19001141-2	离散数学	80	80	0	0		5		40	40					信息院	
	15015430	数字电子技术 A	64	54	0	10		4		64						工学院	
	19001730	计算机组成原理 A	56	48	0	8		3.5		56						信息院	
	19001740	计算机组成原理 A (课程设计)					2	2		(2)						信息院	
	19001270	数据结构 A	64	48	0	16		4		64						信息院	
	19001750	数据结构 A (课程设计)					2	2		(2)						信息院	
	15004430	概率论与数理统计 B	56	52	4	0		3.5		56						理学院	
专业教育平台	专业核心课 (必修)	19001940	Java 程序设计	40	30	0	10		2.5		40					信息院	
		19001950	Java 程序设计 (课程设计)					1	1		(1)					信息院	
		19001760	操作系统 A	48	40	0	8		3		48					信息院	
		19001770	操作系统 A (课程设计)					2	2		(2)					信息院	
		19001840	数据库系统	48	40	0	8		3			48				信息院	
		19001850	数据库系统 (课程设计)					2	2			(2)				信息院	
		19001860	计算机网络 A	48	40	0	8		3			48				信息院	
		19001870	计算机网络 A (课程设计)					2	2			(2)				信息院	
		19001100	计算机体系结构	48	40	0	8		3			48				信息院	
		19001060	大数据技术基础	32	24	0	8		2			32				信息院	
		19001350	探索性数据分析	40	26	0	14		2.5			40				信息院	
		19000980	NoSQL 数据库原理	40	28	0	12		2.5			40				信息院	
		15022280	分布式系统与云计算概论	32	24	0	8		2				32			信息院	
		19001080	机器学习	48	34	0	14		3				48			信息院	
	19000940	Hadoop 集群计算原理及编程	32	24	0	8		2				32			信息院		
	19001250	实时大数据处理技术	32	28	0	4		2					32		信息院		
	19001070	大数据专业实践					3	3						(3)	信息院		
	大数据应用选修模块	19000970	Linux 应用		32	22	0	10		2			32			信息院	
		19001050	大数据采集与集成		32	24	0	8		2			32			信息院	
		16000020	科技论文写作		24	24	0	0		1.5				24		信息院	
		19001480	自然语言处理		32	28	0	4		2			32			信息院	
		19001290	数据可视化		48	24	0	24		3			48			信息院	
		19001470	预测与决策		32	32	0	0		2				32		经管院	
		15007870	计算机网络安全		32	22	0	10		2				32		信息院	
		19001240	时间序列分析 B		32	20	0	12		2				32		理学院	
		19001150	林业大数据分析 with 决策		32	32	0	0		2				32		信息院	
		系统开发选修模块	19001920	C++ 程序设计		40	24	0	16		2.5	40					信息院
			19001010	Web 前端开发		48	28	0	20		3	48					信息院
19001820			Python 应用		32	24	0	8		2		32				信息院	
19001980			软件设计模式 (双语)		40	32	0	8		2.5		40				信息院	
15012620			软件工程 B		40	28	2	10		2.5			40			信息院	
15018730	移动开发技术			40	26	0	14		2.5				40		信息院		

北京林业大学 2020 级本科人才培养方案-信息学院

毕业论文(设计)							8							✓	✓	信息院		
综合 拓展 环节	19001640	军事理论					2		✓							学生处		
	19001650	军事技能					2		✓							学生处		
	15020701-2	志愿服务与公益劳动					2		✓	✓						信息院		
	15002471-4	大学英语自主听说							✓	✓	✓	✓				外语院		
	15002450	大学生素质拓展计划					3		✓	✓	✓	✓	✓	✓		校团委		
	17001000	大学生科技创新							✓	✓	✓	✓	✓	✓	✓	信息院		
	19000010	大学生职业生涯规划								✓						招就处		
	19000020	就业创业指导											✓			招就处		
	17000900	创新创业教育	根据《北京林业大学本科生创新创业学分 管理与应用办法(试行)》执行								✓	✓	✓	✓	✓	✓	教务处	
必修课合计			1968	1642	80	246	17	131	340	488	412	232	288	176	32	0	—	
选修课合计			536	390	2	144	0	33.5	0	88	0	64	72	144	168	0	—	
必修实践环节合计			—	—	—	—	—	48	—	—	—	—	—	—	—	—	—	
毕业生应取得总学分			174						公共选修课学分					7				
									通识必修课学分					37.5				
									暑期学期学分					3				
									学科基础教育平台学分					53				
									专业核心课学分					40.5				
									本专业选修课学分					16				
									毕业论文(设计)学分					8				
									综合拓展环节学分					9				

数据科学与大数据技术专业重点课程简介

大数据技术导论: 本课程是数据科学与大数据专业的专业基础课程,也是 NoSQL 数据库原理、Hadoop 集群编程、分布式系统和云计算等专业课程的重要基础。课程重点围绕大数据数据库技术的生态系统,介绍大数据概述、大数据处理架构 Hadoop、分布式文件系统 HDFS、分布式数据库 HBASE 及大数据处理的编程模式 MapReduce。通过本课程的学习,培养学生大数据思维意识,使学生理解大数据技术的产生背景和技术生态系统,掌握大数据各类技术的基本原理和关系,并能够搭建大数据环境并具备从事简单大数据分析和开发的能力。

Java 程序设计: 本课程是讲述如何使用 Java 语言进行面向对象程序设计的课程,是数据科学与大数据技术的必修课,也是相关后续大数据实践课程的编程基础课。主要内容包括:Java 语言概述、Java 语言基础、类和对象、类的高级特性、常用类、异常处理、输入输出流、图形用户界面、多线程、Java 网络程序设计、Java 数据库操作等。本课程培养学生用 Java 程序解决实际问题的能力,强化面向对象的编程意识,为后续相关课程、将来的工作积累编程经验。

数据库系统: 本课程是离散数学、程序设计基础、数据结构和操作系统等课程之后的一门专业必修课,主要介绍关系型数据库技术的原理、方法和发展趋势。课程采用理论和实践相结合的方法,加强实践环节的教学,结合典型实例和关系数据库管理系统讲解数据库设计、使用和维护的过程。通过本课程的学习,使学生理解数据库的基本概念,掌握基本的数据库技术和方法,为后续的 NOSQL 和数据挖掘等课程打好基础。

NoSQL 数据库原理: 本课程是大数据技术基础和数据库系统等课程之后的一门专业必修课,主要介绍 NoSQL 数据库的原理、关键技术实现方案、应用场景和发展趋势。课程采用理论与实践相结合的方法,通过大数据实训平台,借助典型案例和 NOSQL 数据库实现方案,系统讲解非关系型数据库的架构、设计、开发和操作方法。通过本课程的学习,使学生系统理解 NoSQL 数据库的基本原理,掌握常见 NoSQL 数据库的操作方法,为后续大数据分析和系统开发等工作奠定基础。

分布式系统与云计算概论: 通过本课程的学习,使学生初步了解分布式系统的架构,云计算技术的相关概念以及云计算技术的发展。在介绍云计算相关概念的基础上,进一步介绍分布式云计算程序开发技术,使得学生实际接触并使用云计算技术,激发学生的学习兴趣,深入了解不同云计算产品对问题的适应性,丰富学生对云计算领域的认知,增加分布式系统的开发经验,为今后在相关领域的开发和研究打下坚实的基础。

大数据采集与集成: 本课程是大数据技术基础之后的一门专业必修课,通过本课程的学习,可以让学生掌握大数据的采集与集成的基础知识,包括大数据采集和集成的基本概念和原理、大数据采集和集成框架的工作原理、大数据采集和集成框架的体系结构等。课程学习使学生熟悉和掌握大数据采集与集成的基本概念和实现技术,了解目前大数据采集和集成框架的发展及应用背景,为开展大数据复杂系统的设计和实现以及复杂大数据科学分析和研究工作打下良好的基础。