

附件：

北京大学本科专业（方向）设置申请表

（2022 年修订）

专业（方向）名称：信息与计算科学专业（数据科学方向）

专业代码：070102

所属学科门类及专业类：理学

学位授予门类：本科

修业年限：四年

申请时间：2024 年

专业（方向）负责人：李铁军

联系电话：010-62757592

教务部制

1. 申报专业（方向）基本情况

专业代码	070102	专业（方向）名称	信息与计算科学专业（数据科学方向）
学位	本科	修业年限	四年
专业类	信息与计算科学专业	专业类代码	0701
门类	理学	门类代码	07
所在院系名称	数学科学学院		
学校相近专业情况			
相近专业 1	（填写专业/专业方向名称） 数据科学与大数据技术	（开设年份） 2016	该专业教师队伍情况 （见3.2 教师基本情况表）
相近专业 2	信息与计算科学	1913	该专业教师队伍情况 （同上）
相近专业 3	统计学	1913	该专业教师队伍情况 （同上）

2. 申报专业（方向）人才需求情况

申报专业（方向）主要就业领域	互联网和科技公司、金融行业、健康医疗领域、零售和电商行业、制造业和物流领域、政府部门和非营利组织、咨询公司、市场调研机构、科研院所等	
据国家网信办发布的数据显示，2017 年到 2021 年，中国数据产量从 2.3ZB 增长至 6.6ZB，四年间翻了将近 3 倍。2021 年中国数字经济规模已经突破 45 万亿元，占 GDP 的 39.8%；2022 年中国大数据产业规模已达 1.57 万亿元，成为推动数字经济发展的主要力量。据中国社会科学院测算，“十四五”时期，中国数字经济无论是数字产业化还是产业数字化都将延续快速增长势头，数字经济整体的复合年均增长率为 11.3%；到 2025 年，数字经济增加值规模有望达到 32.67 万亿元。数字经济的高速发展，导致大数据人才的培养数量和速度远远达不到产业规模的增速，专业人员奇缺。据中国商业联合会数据分析专业委员会资料显示，目前中国大数据从业人员只有约 30 万人，人才缺口高达 120 万人，到 2025 年缺口将达到 200 万人。 北京大学在数据科学的教学方面强调基础和应用的结合，鼓励学生在数学模型，算法和计算机等方面实现全面发展。学生应该具备扎实的理论基础，同时也具备扎实的实践能力。以有效地避免目前这种数学专业学生缺乏实践能力，计算机专业的学生缺乏理论功底的尴尬局面。这样的人才将会得到信息产业的广泛欢迎。 总之，大数据给科学和教育事业的发展提供了前所未有的机会，同时也提出了前所未有的挑战。它将对现有的科研和教学体制带来大幅度的变革，对科学与产业之间的关系、科学与社会之间的关系带来大幅度的变革。 首先是数据科学将成为科研体系中的重要部分，并逐渐达到与包括物理、化学、生命科学等学科在内的自然科学分庭抗礼的地位。未来的科研和教育体制应该由两条主线组成：一条是以基本原理为主线。现在的物理学、化学、机械工程等学科，以及生命科学、材料科学、天体物理、地球科学等学科的大部分都是沿着这样一条主线展开的。另一条是以数据为主线。它包括统计学、数据挖掘和机器学习、生物信息学、天体信息学、以及许多社会科学的学科。它还包括一些新兴的学科，如计算广告学。数据科学的兴起，将极大地推动许多社会科学学科朝着量化的方向发展，使他们逐步由经验性的模式转变成科学性的模式。 其次是科学研究和市场、和产业的联系将变得更加密切，从发现基本原理到产业化的周期将会被大大地缩短。再次，数据的主要来源之一是社会，如互联网、社交网络、公共交通、智慧城市等等。所以数据科学的研究与我们的日常生活、与社会有着密切的联系。科学研究最重要的一环是提出前瞻性的问题。提不出问题，就只能跟在别人后面，走一条从文献到文献的路子。对我国的科技界来讲，在很多学科，由于来自实际应用领域的限制，提出前瞻性问题确实是件很困难的事情。但数据科学则不然。由于我国人口众多这一特殊情况，和我们特殊的文化、文字、历史背景和社会发展的需要，我们在数据科学领域的很多问题自然就是前瞻性的。关键是我们能否用前瞻性的方法去面对这些问题。如果做好了这一点，我们在数据科学领域就自然而然地走到了世界的前沿。 综上所述，我们认为“数据科学”方向既有迫切的社会需要，又有直接的教育和研究意义，因此我们将依托北京大学数学科学学院，申请设置信息与计算科学专业（数据科学方向）。我们相信这个方向的设立，将极大促进我国高素质人才的培养，提高我校本科生教育的整体水平，并进一步促进数学，计算机科学以及广泛的基础和传统学科的发展，为我国创立世界一流的研究型大学提供重要支撑，也为中国本科高等教育在跨学科人才培养方面探索新路，积累经验。		
申报专业（方向）人才需求调研情况 （可上传合作办学协议等）	年度计划招生人数	15
	预计升学人数	10

	预计就业人数	5
	其中：华为技术有限公司	2
	百度在线网络技术（北京）有限公司	1
	旷视科技有限公司	1
	北京三快在线科技有限公司	1

3. 教师及课程基本情况表

3.1 教师及开课情况汇总表

专任教师总数	45
具有教授（含其他正高级）职称教师数及比例	20，44%
具有副教授及以上（含其他副高级）职称教师数及比例	8，18%
具有硕士及以上学位教师数及比例	45，100%
具有博士学位教师数及比例	45，100%
35 岁及以下青年教师数及比例	4，8%
36-55 岁教师数及比例	26，58%
兼职/专职教师比例	无兼职
专业核心课程门数	7
专业核心课程任课教师数	17

3.2 教师基本情况表

姓名	性别	出生年月	拟授课程	专职/兼职	专业技术职务	学历	最后学历毕业学校	最后学历毕业专业	最后学历毕业学位	研究领域
鄂维南	男	1963/09	机器学习与科学计算、现代应用数学介绍	专职	讲席教授	博士	美国加州大学洛杉矶分校	数学	理学博士	机器学习的基础算法和基础理论；计算科学中基于机器学习的新算法；高维控制问题的基于机器学习的算法和应用
李铁军	男	1974/10	数值分析、随机模拟方法	专职	教授	博士	北京大学	数学	理学博士	随机模型及算法
董彬	男	1981/07	研究型学习、机器学习与科学计算、图像处理中的数学方法	专职	教授	博士	美国加利福尼亚大学洛杉矶分校	数学	理学博士	科学计算、机器学习及其在计算成像和数据分析中的应用。
胡俊	男	1971/09	应用数学导论、数值代数	专职	教授	博士	中国科学院数学与系统科学研究院	数学	理学博士	非标准有限元方法的研究
李若	男	1973/07	高等数学 C（一）、	专职	教授	博士	北京大学	数学	理学博士	偏微分方程数值解；计算流体力学

			流体力学引论							和数学软件
张磊	男	1978/12	计算方法B、计算系统生物学	专职	教授	博士	美国宾州州立大学	数学	理学博士	计算和应用数学，包括稀有事件与解景观的算法与应用，计算材料科学，计算系统生物学，机器学习等
杨超	男	1979/1	并行计算	专职	教授	博士	中国科学院软件所	数学	理学博士	高性能科学与工程计算
邵嗣烘	男	1982/10	数学模型、高维数值方法	专职	教授	博士	北京大学	数学	理学博士	量子化学和量子物理中的算法、微分方程数值解、图谱理论和谱聚类算法
文再文	男	1982/3	大数据分析中的算法、最优化方法	专职	教授	博士	美国哥伦比亚大学	运筹学	理学博士	最优化算法与理论、机器学习、数据科学
吴磊	男	1990/3	高等数学B(二)、机器学习数学导引	专职	研究员	博士	北京大学	数学	理学博士	深度学习数学理论
黄得	男	1992/6	高等数学B(一)、(二)	专职	研究员	博士	美国加州理工学院	数学	理学博士	随机矩阵、集中不等式、数值偏微分方程、流体力学方程
黄政宇	男	1992/3	反问题与数据同化	专职	助理教授	博士	美国斯坦福大学	数学	理学博士	计算数学、计算工程、数据科学
姜明	男	1963/9	计算机图形学、现代信息处理选讲	专职	教授	博士	北京大学	数学	理学博士	图像重建和图像处理
林作铨	男	1963/3	人工智能、数理逻辑	专职	教授	博士	北京航空航天大学	数学	理学博士	人工智能、计算机软件、应用逻辑
马尽文	男	1962/07	信息科学基础、智能计算	专职	教授	博士	南开大学	数学	理学博士	神经计算；模式识别；图像处理；大数据挖掘；生物信息学

牟克典	男	1975/09	机器学习基础、机器学习、大数据分析机器学习基础	专职	副教授	博士	北京大学	数学	理学博士	人工智能
孙猛	男	1978/12	程序设计技术与方法、程序设计语言原理、论文写作指导	专职	教授	博士	北京大学	数学	理学博士	软件理论、形式化方法论；模型检查；定理证明；协调模型和语言；余代数理论；信息物理系统；区块链与智能合约；大数据分析；深度学习
夏壁灿	男	1968/09	理论计算机科学基础	专职	教授	博士	四川大学	数学	理学博士	符号计算；机器证明；自动推理；程序验证

3.3 专业（方向）核心课程表（以下表格数据由申报专业（方向）填写）

课程名称	课程 总学时	课程 周学时	拟授课教师	授课学期
数学分析 III	80	5	王崑	二上（二年级秋季学期）
几何学	96	6	包志强	一上（一年级秋季学期）
抽象代数	48	3	徐茂智	二上（二年级秋季学期）
复变函数	48	3	郭帅	二下（二年级春季学期）
常微分方程	48	3	柳彬	二下（二年级春季学期）
概率论	48	3	章复熹	二下（二年级春季学期）
机器学习基础	48	3	牟克典	二下（二年级春季学期）

4. 专业（方向）主要带头人简介

姓名	鄂维南	性别	男	专业技术职务	中国科学院院士、北京大学讲席教授	行政职务	北京大数据研究院院长、北京大学国际机器学习研究中心主任
现在所在单位	北京大学数学科学学院 北京大数据研究院 北京大学国际机器学习研究中心			拟承担课程	机器学习与科学计算、现代应用数学介绍		
最后学历毕业时间、学校、专业		1989 年，美国加州大学洛杉矶分校（UCLA），数学系，博士					
主要研究方向		机器学习的基础算法和基础理论；计算科学中基于机器学习的新算法；高维控制问题的基于机器学习的算法和应用					
从事教育教学改革研究及获奖情况（含教改项目、研究论文、慕课、教材等）							
从事科学研究及获奖情况		国家重点研发计划，数学与应用研究专项，2022YFA1004300, 药物靶标结合强度和结合过程的可计算建模及高效算法，2022.12 至 2027.11，150 万元，课题负责人；国家自然科学基金委员会，重大研究计划，可解释、可通用的下一代人工智能方法战略研究项目，92270001，人工智能未来发展趋势战略研究，2023.01-2025.12，300 万元，项目负责人。 2003 年获国际工业与应用数学协会 Collatz 奖；2009 年获美国工业与应用数学学会 Kleinman 奖；2014 年获美国工业与应用数学学会 von Karman 奖；2019 年获得美国工业与应用数学学会和瑞士 ETH 共同颁发的 Peter Henrici 奖；2020 年获 ACM Gordon Bell 奖；2023 年获国际工业与应用数学协会 Maxwell 奖。 曾受邀在包括国际数学家大会（ICM）、国际工业与应用数学大会（ICIAM）、美国数学会年会（AMS）、美国物理学会年会（APS）、美国化学学会年会（ACS）、美国化工协会年会（AIChE）、美国理论化学大会（ACTC）、国际计算力学大会（WCCM）等多个数学、应用数学、物理、化学、力学等领域的顶级会议上作报告。2022 年国际数学家大会 1 小时报告人。2022 年国际机器学习大会特邀报告人。					
近三年获得教学 研究经费（万元）				近三年获得科学 研究经费（万元）		485 万元	
近三年给本科生授课课程及学时数		《现代应用数学介绍》、《机器学习与科学计算》、《应用数学讨论班》，80 学时		近三年指导本科毕业设计（人次）		0	

注：填写三至五人，只填本专业专任教师，每人一表。

4. 专业（方向）主要带头人简介

姓名	李铁军	性别	男	专业技术职务	教授	行政职务	数学科学学院信息与科学计算系主任、大数据科学研究中心副主任
现在所在单位	北京大学数学科学学院 北京大学大数据科学研究中心			拟承担课程	数值分析、随机模拟方法		
最后学历毕业时间、学校、专业		2001 年，北京大学数学院，数学，博士					
主要研究方向		随机模型及算法					
从事教育教学改革研究及获奖情况（含教改项目、研究论文、慕课、教材等）							
从事科学研究及获奖情况		国家自然科学基金委员会，国家杰出青年科学基金项目，11825102，随机模型及算法，2019.01-2023.12，245 万元，项目负责人；国家重点研发计划，数学与应用研究专项，2021YFA1003300，5.5G 大规模 MIMO 通信系统的超分辨率参数估计和补全问题，2022.05 至 2026.04，1200 万元，项目负责人。 在包括 Nature Comm.，Proc.Nat.Acad.Sci.，Phys Rev.X，Comm.Math. Phys.，Briefings Bioinformatics, PLoS Comp.Biol.，Multi.Model. Simul.，J.Chem.Phys.，Bioinformatics 等在内的国际重要学术期刊上发表论文 60 余篇。2010 年入选教育部“新世纪人才支持计划”、2012 年获国家自然科学基金委优秀青年基金项目、2018 年获国家自然科学基金委杰出青年基金项目。					
近三年获得教学研究经费（万元）				近三年获得科学研究经费（万元）		324 万	
近三年给本科生授课课程及学时数		《数值分析》、《随机模拟方法》、《现代应用数学介绍》，112 学时		近三年指导本科毕业设计（人次）		5	

注：填写三至五人，只填本专业专任教师，每人一表。

4. 专业（方向）主要带头人简介

姓名	董彬	性别	男	专业技术职务	教授	行政职务	国际机器学习研究中心副主任
现在所在单位	北京国际数学研究中心、国际机器学习研究中心、大数据分析与应用国家工程实验室、国家生物医学成像科学中心			拟承担课程	研究型学习、机器学习与科学计算、图像处理中的数学方法		
最后学历毕业时间、学校、专业		2009 年，美国加州大学洛杉矶分校，数学系，博士					
主要研究方向		计算成像和图像处理中的数学问题；机器学习与科学计算					
从事教育教学改革研究及获奖情况（含教改项目、研究论文、慕课、教材等）		2019 年北京高等教育“本科教学改革创新项目”数学基础学科拔尖人才培养创新项目研究； 2021 年教育部“基础学科拔尖学生培养计划 2.0”优秀教师奖					
从事科学研究及获奖情况		第二期“新基石研究员”项目，2024 年至 2031 年，1620 万，在研，主持；北京市科学技术委员会，中央引导地方专项项目，基于多模态大模型的消化系统肿瘤诊疗技术研究及示范应用，2023-09 至 2025-09，600 万，在研，参与 2014 年获得求是杰出青年学者奖， 2015 年入选中组部“千人计划”青年项目，2019 年入选科技部创新人才推进计划，2022 年受邀在世界数学家大会（ICM）做 45 分钟报告，2023 年第二期“新基石研究员”项目。					
近三年获得教学研究经费（万元）				近三年获得科学研究经费（万元）		1860 万	
近三年给本科生授课课程及学时数		《研究型学习》、《机器学习与科学计算》、《现代应用数学介绍》，184 学时		近三年指导本科毕业设计（人次）		9	

注：填写三至五人，只填本专业专任教师，每人一表。

5. 教学条件情况表

可用于该专业（方向）的教学实验设备总价值（万元）	300 万以上	可用于该专业（方向）的教学实验设备数量（千元以上）	279
开办经费及来源	年平均教学经费约 400 万元，来源于国拨经费		
生均年教学日常支出（元）	5000 元		
实践教学基地（个）	14 个（郑州数字创新中心教育基地，甘肃酒泉教育基地，河北省雄安新区北沙口乡教育基地，黑龙江佳木斯教育基地，辽宁盘锦教育基地，陕西华阴教育基地，浙江龙港姜立夫故居教育基地，中国归谷嘉善科技园教育基地，中国空气动力研究与发展中心超高速空气动力研究所教育基地，重庆大数据研究院教育基地，山西榆次教育基地，北大计算与数字经济研究院基地，山西晋城教育基地，安徽亳州教育基地）		
教学条件建设规划及保障措施	学院本科被教育部遴选为国家“理科基础科学研究和科学人才培养基地”、数学学科拔尖人才培养基地”。学院拥有“数学及其应用”教育部重点实验室、“统计与信息技术”教育部-微软重点实验室等教育部重点实验室；还拥有北京大学科学与工程计算中心、北京大学统计科学中心等多个中心。学院学科门类齐全，教学与科研并重，理论与应用并举，携手北京国际数学研究中心，已成为具有重要国际影响的数学科学研究和人才培养基地。 北大数学学院拥有一支实力雄厚、学风严谨的师资队伍。学院现有教师 101 人，其中教授 59 人、副教授 27 人、讲师/助理教授 15 人。其中中科院院士 9 名、第三世界科学院院士 4 名、长江特聘教授和长江学者 15 名、国家杰出青年基金获得者 28 名、优秀青年科学基金获得者 11 名、千人计划 10 人、青年千人计划 17 人、中组部“青年拔尖人才”入选者 4 名。 学院拥有五个本科专业。全院的教学工作由负责教学的副院长统一主抓，一流的教学管理人员为全员做好细致专业的教学保障工作。 学院是北京大学较早将计算机和互联网引入到教学和科研的院系。学院设有 4 个实验室。现有高性能集群 2 个，GPU 服务器 16 个，工作站 26 台，微机 220 余台。学院还设有专门的本科生机房（70 个机位）和研究生机房（20 个机位），为学生提供良好的上机服务。 数学学院图书馆是北大图书馆的一个分馆，藏书量丰富。数学学院还设有本科生阅览室。为本科生提供了充足的学习及课外用书。		

主要教学实验设备情况表

教学实验设备名称	型号规格	数量（台/件）	购入时间	设备价值（千元）
服务器	IW4200-8G	2	2016-03-09	123
服务器	PowerEdge T630	2	2016-12-01	108.33

服务器	NF5280M5	4	2018-03-08	205
服务器	R720	1	2013-08-16	107.4
服务器	X3500M5	2	2016-12-07	99.85
服务器	NF5280M4	2	2017-05-08	90.3
磁盘阵列	DELL MD1200	1	2013-10-08	52
工作站	SCW4150	2	20150702	61.8
工作站	HP Z8 G4 Workstation	2	20191016	58
服务器	NX5440M4	10	20180308	36.9
服务器	NF5280M4	4	20180308	27.6
服务器	P80005U	2	20161215	38.88
服务器	IR2200	1	20160503	15.3
工作站	Mac mini M1	1	20211125	6.599

6. 申请增设专业（方向）的理由和基础

应包括申请增设专业（方向）的主要理由、支撑该专业（方向）发展的学科基础、学校专业发展规划等方面的内容）（如需要可加页）

一、数据科学方向是大数据时代的现实要求和传统学科发展的动力需求

数据科学是一门新兴的学科。它和已有的信息科学、统计学、机器学习等学科不同。

作为一门学科，数据科学所依赖的两个因素是数据的广泛性和多样性，以及数据研究的共性。现代社会的各行各业都充满了数据。而且这些数据也是多种多样，不仅包括传统的结构型数据，也包括网页、文本、图像、视频、语音等非结构型数据。正如我们后面将要讨论到的，数据分析本质上都是在解反问题，而且常常是随机模型的反问题。所以对它们的研究有着很多的共性。

数据科学主要包括两个方面：用数据的方法来研究科学和用科学的方法来研究数据。前者包括生物信息学、天体信息学、数字地球等领域。后者包括统计学、机器学习、数据挖掘、数据库等领域。这些学科都是数据科学的重要组成部分。但只有把它们有机地放在一起，才能形成整个数据科学的全貌。

数据科学对学科发展提供了前所未有的机遇和挑战。要充分利用好这个机会，我们就必须建立起一套新的科学和教育体系。在大学的层面，要赋予数据科学其应有的地位，建立起跨学科，全方位的数据科学研究平台；进一步完善和企业合作创新的机制；培养适应学术界和企业界需求的数据科学人才。

数据科学也将对许多传统学科的发展带来极大的影响。首先是对数学。数学的发展主要来自两个方面的推动力：一是来自数学内部，学科自身的完善带来的推动；二是来自外部，由其它学科，社会或工业发展的需要而带来的推动。就目前的现状而言，第一方面的推动力对数学的影响要远远超过第二方面的推动力。这样造成的结果是，一方面，数学作为一门学科，其重要性已经得到广泛的认可。而另一方面，数学家作为一个群体，其对社会和科学整体发展的影响却难以得到承认。数据科学的基本原理本身就来自于数学。所以数据科学在数学和实际应用之间建立起了一个直接的桥梁。而这些实际应用正是来自于信息服务等现代产业中最为活跃的一部分。这对数学来说，实在是一个千载难逢的机会。数据的分析几乎涉及到了现代数学的所有分支。甚至于像表示论这样的极其抽象的分支，在数据的领域也有其发挥作用的余地。所以数据科学对数学的要求和推动是全面的，而不是仅仅局限在几个领域。数据应该成为数、图形和方程之外数学研究的基本对象之一。计算数学中的很多基础工具，例如最优化计算、概率论、随机过程、动力系统，等，瞬间被迁移到数据科学相关的算法中，得以利用。与此同时，数据科学相关计算方法在实际应用过程中，也推动了计算数学的发展。

数据科学对计算机科学的发展也会带来很大的影响。图灵奖得主 John Hopcroft 曾经指出，在过去的几十年里，计算机科学的研究对象主要是计算机本身，包括硬件和软件。以后计算机科学的发展将主要围绕着应用展开。而从计算机科学自身来看，这些应用领域提供的主要研究对象就是数据。虽然计算机科学一贯重视数据的研究，但数据在其中的地位将会得到更进一步的加强。在当前数据爆炸和信息技术对社会、经济生活等各方面渗透的大背景下，利用前沿的数据科学和信息技术，探索自然科学、工程、经济、社会等多学科的交叉领域已成必然趋势。信息技术为大数据的收集、处理、分析和应用提供了基础设施和工具，同时大数据也驱动了信息技术的发展和创新。这两者相互支持、相互促进，共同推动着数据驱动的决策和创新的发展。发展数据科学既能更加适应未来的需要，又能尽快地把应用数学、计算数学、计算机科学和信息科学等学科中的有生力量调动起来以开展工作。

二、北京大学数学科学学院具有设置“信息与计算科学专业（数据科学方向）”的优越条件

北大的数据科学研究有着良好的基础。统计中心、数学学院和信息科学技术学院在基础研究方面的实力在国内处于领先地位。信息科学技术学院，中国社会科学调查中心，生命科学学院，医学信息中心等单位在应用方面的工作也处于领先地位。随着当代信息科学迈入数据科学时代，学科的发展迫切呼唤信息科学与科学计算的融合，北京大学数学学院信息与计算科学系由此应运而生。计算数学方向：开设的课程主要包括应用数学导论、数值代数、数值分析、数值偏微分方程、最优化方法、并行计算、流体力学引论、随机模拟方法、数学物理反问题、生物数学导论、多尺度建模及计算、图像处理的数据理论、数据中的数学、大数据分析中的算法等。信息科学方向：与欧洲、美国、澳大利亚、日本、新加坡等地的多所大学及研究所有密切的学术合作与交流关系，科研能力突出的本科生及研究生可被推荐出国进行交流或攻读博士学位。毕业生可在科研机构或高校继续从事信息科学相关的科研工作，也可在生物、金融、交通、医疗等自然科学和社会科学领域或业界从事大数据的采集、管理、分析与处理方面的工作。北京大学信息与计算科学专业（数据科学方向）由北京大学数学学院信息与计算科学系从事数据研究的优秀教师参与课程设计与教学，其中包括多名中国科学院院士、国家千人计划专家、长江学者、国家杰出青年基金获得者，实现教学资源的最优配备。2015年，经教育部批准，北京大学数学科学学院在国内高校之中率先设立数据科学与大数据技术本科专业。很多基础课程已经有很好的资源，新兴专业课程也已经具备相当的研究基础。在课程设计与教学中强调对科学深入而融会贯通的理解，培养学生的主动性、批判性、全方位、跨学科的创新型思维。

北京大学 1) 学科综合优势：北京大学学科种类齐全，综合优势明显，共有国家一级重点学科 18 个，国家二级重点学科 25 个。在大数据交叉学科建设方面的优势学科众多，如：数学、统计学、计算机、软件；应用经济学、社会学、新闻传播学、心理学；基础医学、临床医学、口腔医学、公共卫生与预防医学等。2) 交叉学科平台：北京大学前沿交叉学科研究院是跨学科的研究平台，研究院开展了一系列与生命、医学、信息和材料等先进科学与高新技术相关的创新性研究工作；具有良好学术交流环境、优良学科交叉研究条件；积累了跨学科研究生培养的经验。3) 校企合作基础：与百度、腾讯、联想、方正、中移动、联通等知名企业开展了大数据相关的研究和人才培养合作，这些企业可为科学研究和人才培养提供真实的“大数据”和大数据运行环境。以上都为数据科学方向的设立奠定了基础。

三、人才需求预测

据国家网信办发布的数据显示，2017 年到 2021 年，中国数据产量从 2.3ZB 增长至 6.6ZB，四年间翻了将近 3 倍。2021 年中国数字经济规模已经突破 45 万亿元，占 GDP 的 39.8%；2022 年中国大数据产业规模已达 1.57 万亿元，成为推动数字经济发展的力量。据中国社会科学院测算，“十四五”时期，中国数字经济无论是数字产业化还是产业数字化都将延续快速增长势头，数字经济整体的复合年均增长率为 11.3%；到 2025 年，数字经济增加值规模有望达到 32.67 万亿元。数字经济的高速发展，导致大数据人才的培养数量和速度远远达不到产业规模的增速，专业人员奇缺。据中国商业联合会数据分析专业委员会资料显示，目前中国大数据从业人员只有约 30 万人，人才缺口高达 120 万人，到 2025 年缺口将达到 200 万人。

北京大学在数据科学的教学方面强调基础和应用的结合，鼓励学生在数学模型，算法和计算机实现等方面全面发展。学生应该具备扎实的理论基础，同时也具备扎实的实践能力。以有效地避免目前这种数学专业学生缺乏实践能力，计算机专业的学生缺乏理论功底的尴尬局面。这样的人才将会得到信息产业的广泛欢迎。

总之，大数据给科学和教育事业的发展提供了前所未有的机会，同时也提出了前所未有的挑战。它将对现有的科研和教学体制带来大幅度的变革，对科学与产业之间的关系、科学与社会之间的关系带来大幅度的变革。

首先是数据科学将成为科研体系中的重要部分，并逐渐达到与包括物理、化学、生命科学等学科在内的自然科学分庭抗礼的地位。未来的科研和教育体制应该由两条主线

组成：一条是以基本原理为主线。现在的物理学、化学、机械工程等学科，以及生命科学、材料科学、天体物理、地球科学等学科的大部分都是沿着这样一条主线展开的。另一条是以数据为主线。它包括统计学、数据挖掘和机器学习、生物信息学、天体信息学、以及许多社会科学的学科。它还包括一些新兴的学科，如计算广告学。数据科学的兴起，将极大地推动许多社会科学学科朝着量化的方向发展，使他们逐步由经验性的模式转变成科学性的模式。

其次是科学研究和市场、和产业的联系将变得更加密切，从发现基本原理到产业化的周期将会被大大地缩短。再次，数据的主要来源之一是社会，如互联网、社交网络、公共交通、智慧城市等等。所以数据科学的研究与我们的日常生活、与社会有着密切的联系。

科学研究最重要的一环是提出前瞻性的问题。提不出问题，就只能跟在别人后面，走一条从文献到文献的路子。对我国的科技界来讲，在很多学科，由于来自实际应用领域的限制，提出前瞻性问题确实是件很困难的事情。但数据科学则不然。由于我国人口众多这一特殊情况，和我们特殊的文化、文字、历史背景和社会发展的需要，我们在数据科学领域的很多问题自然就是前瞻性的。关键是我们能否用前瞻性的方法去面对这些问题。如果做好了这一点，我们在数据科学领域就自然而然地走到了世界的前沿。

综上所述，我们认为“数据科学”方向既有迫切的社会需要，又有直接的教育和研究意义，因此我们郑重地向学校建议，通过依托北京大学数学科学学院，充分利用“信息与科学计算系”的优势资源，按照新的教学设计，在我院增设“信息与计算科学专业（数据科学方向）”。我们相信这个方向的设立，将极大促进我国大数据领域高素质人才的培养，加快填补中国大数据人才缺口，并进一步促进数学、计算机科学、信息科学以及广泛的基础和传统学科的发展，为我国创立世界一流的研究型大学提供重要支撑，也为中国本科高等教育在跨学科人才培养方面探索新路，积累经验。

7. 申请增设专业（方向）人才培养方案

包括培养目标、基本要求、修业年限、授予学位、主要课程、主要实践性教学环节和主要专业实验、教学计划等内容）（如需要可加页）

一、专业简介

数据科学综合运用统计学、计算机科学、应用数学等学科提供的现代数据分析工具和方法从数据中自动寻找规律或者有价值的信息。具体地，它是运用概率统计、并行与分布式计算、人工智能、机器学习等综合知识研究来自工业、生物医药、金融证券和社交网络等众多领域的较大规模或结构复杂数据集的高效采集、高效存储、高效管理、精确建模、深入分析和精准预测的新兴交叉学科。

2015 年，经教育部批准，北京大学数学科学学院在国内高校之中率先设立数据科学与大数据技术本科专业。2017 年首批该专业本科生毕业。

本专业的教师来自数学科学学院信息与计算科学系和概率统计系。

二、培养目标

本专业致力于培养掌握数学、计算机、统计等数据科学相关领域基础理论知识，以及数据建模、机器学习、并行与分布式计算、统计推断等方法和技术，从事数据建模、数据分析与挖掘算法等问题的研究和大数据系统开发的研究型和技术型人才。毕业生可在科研机构或高校继续深造、从事数据科学相关的科研工作，也可在生物、金融、交通、医疗等自然科学和社会科学领域或业界从事大数据的采集、管理、分析与处理方面的工作。

三、培养要求

通过四年的学习，学生应扎实地掌握数学、统计、信息科学的基础知识和数据建模、数据分析与挖掘等基本技能，英语水平达到国家四级，具备独立学习的能力、初步的研究能力以及较强的适应不同社会职业需要的能力。

四、毕业要求及授予学位类型

学生在学校规定的学习年限内，修完培养方案规定的内容，成绩合格，达到学校毕业要求的，准予毕业，学校颁发毕业证书；符合学士学位授予条件的，授予学士学位。

授予学位类型：理学学士学位毕业总学分：135-141 学分

具体毕业要求包括：

1、公共基础课程：42-48 学分	1-1 公共必修课：30-36 学分
	1-2 通识教育课：12 学分
2、专业必修课程：49 学分	2-1 专业基础课：19 学分
	2-2 专业核心课：24 学分
	2-3 毕业论文（设计）：6 学分
	2-4 其他非课程必修要求：0 学分
3、选修课程：44 学分	3-1 专业选修课：21 学分
	3-2 自主选修课：23 学分

五、课程设置

1、公共基础课程：42-48 学分

1-1 公共必修课：30-36 学分

课号	课程名称	学分	周学时	实践总学时	选课学期及说明
----	------	----	-----	-------	---------

——	大学英语	2-8	——	——	按大学英语教研室要求选课
04031650	思想道德修养与法律基础	2 学分	2 学时		按学校要求选课
04031660	中国近现代史纲要	2 学分	2 学时		按学校要求选课
04031731	毛泽东思想和中国特色社会主义理论体系概论	3 学分	3 学时		按学校要求选课
04031740	马克思主义基本原理概论	3 学分	3 学时		按学校要求选课
04031761	习近平新时代中国特色社会主义思想概论	2 学分	2 学时		按学校要求选课
04031751	形势与政策	2 学分	2 学时		按学校要求选课
61130030	思想政治实践（上） 两个模块任选其一	1 学分			按学校要求选课
61130040	思想政治实践（下）	1 学分			按学校要求选课
04831410	计算概论 B	3	3	0	一上 面向理科院系。学生选“计算概论 B”课程后，需要另选该课程的上机课“计算概

					论 B 上机”。
04831650	计算概论 B 上机	0	2	32	一上 面向理科院系。学生选“计算概论 B”课程后，需要另选该课程的上机课“计算概论 B 上机”。
04831420	数据结构与算法 B	3	3	0	一下 说明：面向理科院系。院系可以根据学科特点选择是否作为必修课程。学生选“数据结构与算法 B”课程后，需要另选该课程的上机课“数据结构与算法上机”。
04830494	数据结构与算法上机	0	2	32	一下 说明：面向理科院系。学生选“数据结构与算法 B”课程后，需要另选该课程的上机课“数据结构与算法上机”。
60730020	军事理论	2	2	0	一上或一下
——	体育系列课程	1*4	2	0	全年

1-2 通识教育课程及学分要求

通识教育课程分为四个系列： I. 人类文明及其传统、 II. 现代社会及其问题、 III. 艺术与人文、 IV. 数学、自然与技术，每个系列均包含通识教育核心课、 通选课两部分课程，具体课程列表详见《北京大学本科生选课手册》。

通识教育课程修读总学分为 12 学分。具体要求包括：

（1）至少修读 1 门“通识教育核心课程”（任一系列），且在四个课程系列中每个系列至少修

读 2 学分（通识教育核心课或通选课均可）；

（2）原则上不允许以专业课替代通识教育课程学分；

（3）本院系开设的通识教育课程不计入学生毕业所需的通识教育课程学分；

（4）建议合理分配修读时间，每学期修读 1 门课程。

2. 专业必修课程： 49 学分

2-1 专业基础课： 19 学分

课号	课程名称	学分	周学时	实践总学时	选课学期
----	------	----	-----	-------	------

00132301	数学分析 I	5	6		一上（一年级秋季学期）
00132302	数学分析 II	5	6		一下（一年级春季学期）
00132321	高等代数 I	5	6		一上（一年级秋季学期）
00132323	高等代数 II	4	5		一下（一年级春季学期）

2-2 专业核心课： 24 学分

课号	课程名称	学分	周学时	实践总学时	选课学期
00132304	数学分析 III	4	5		二上（二年级秋季学期）
00132341	几何学	5	6		一上（一年级秋季学期）
00135450	抽象代数	3	3		二上（二年级秋季学期）
00132320	复变函数	3	3		二下（二年级春季学期）
00132340	常微分方程	3	3		二下（二年级春季学期）
00131300	概率论	3	3		二下（二年级春季学期）
00137170	机器学习基础	3	3		二下（二年级春季学期）

注：数学分析 I、 II、 III，高等代数 I、 II，几何学，概率论都同时开设常规班和实验班，均可作为毕业学分，但一种课程班型已修读及格后，不能再修读另一种班型。因课号、 班型不同， 计算学分、 GPA 时，一种班型的及格成绩不能覆盖另一种班型的不及格成绩。

2-3 毕业论文： 6 学分

2-4 其他非课程必修要求： 0 学分

3. 选修课程： 44 学分

3-1 专业选修课： 21 学分

3-1-1 专业必选： 6 学分

课号	课程名称	学分	周学时	实践总学时	选课学期
00130280	计算方法 B	3	3		三上（三年级秋季学期）
00137150	并行与分布式计算基础	3	3		三上（三年级秋季学期）

3-1-2 专业限选： 15 学分

课号	课程名称	学分	周学时	实践总学时	选课学期
00132370	实变函数	3	3		秋季
00132310	微分几何	3	3		秋季
00130161	拓扑学	3	3		秋季
00135460	数理统计	3	3		秋季
00132330	偏微分方程	3	3		秋季
00133010	测度论	3	3		春季
00136660	凸优化	3	3		秋季
00137130	深度学习：算法与应用	3	3		春季
08408010	强化学习：理论与算法	3	3		春季
04630790	数据科学导引	3	3		秋季
00110950	人工智能	3	3		春季
00136720	大数据分析中的算法	3	3		春季
00133050	应用多元统计分析	3	3		秋季
00133070/ 00131100	应用时间序列分析/金融时间序列分析	3	3		秋季
00133110	应用回归分析	3	3		春季
00130730	数理逻辑	3	3		秋季
00110060	算法设计与分析	3	3		秋季
00135290	集合论与图论	3	3		春季
00135040	程序设计技术与方法	3	3		秋季
04831750	程序设计实习	3	4		春季
04830220	数据库概论	3	3		全年

3-2 自主选修课： 23 学分

3-2-1 理学部及信息与工程科学部课程： 12 学分

可以选自理学部及信息与工程科学部中的任何院系，包括数学学院和信息科学技术学院。要求是该院系的专业必修、专业限选或专业任选，不能是通选和公选。

除专业限选课外，以下课程可以作为自主选修课程参考：

课号	课程名称	学分	周学时	实践总学时	选课学期
04831780	自然语言处理导论	2	2		春季
04833040	计算机系统导论	5	4		秋季
00102892	统计学习	3	3		秋季
00102893	生物统计	3	3		春季
00135590/04831300	计算机图象处理/图像处理	3	3		春季
00130830	数字信号处理	3	3		秋季
00130210/ 04830230	计算机图形学	3	3		春季
04831290/ 00110960	模式识别导论/模式识别	3	3		秋季
04834210	计算机网络	4	5		全年
04830270	程序设计语言概论	2	2		全年
04831890	现代信息检索导论	2	2		秋
04831280	可视化与可视计算概论	2	2		春

3-2-2 理学部的非数学学院课程 8 学分，其中要求物理类课程 4 学分。

8 学分全部选普物 I、II 也行，也可以选其他物理课，非物理类课程 4 学分要求是该院系的专业必修、专业限选或专业任选，不能是通选和公选（大学化学和普通生物学除外，普通生物学 A、B、C 只能选其一修）。

3-2-3 在全校课程中选择其余 3 学分。

全校任何课程均可，包括通选和公选。

六、其他

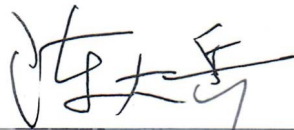
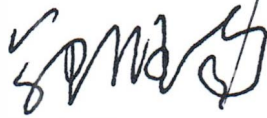
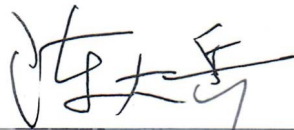
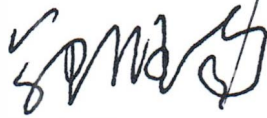
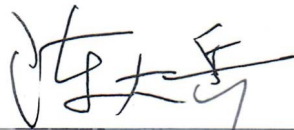
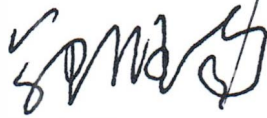
1、保送研究生要求

专业基础课、专业核心课、专业选修课五门（计算方法 B、并行与分布式计算基础、及其他任选三门）进行简单求和作为保研专业排名的依据。

2、上述专业选修 3-1 和学部限选 3-2-1 课程，原则上均以所列课号和课程名称为准。如学生在其他院系选修同名或相似课程原则上不能计入上述两类课程毕业学分。

8. 校内专业（方向）设置评议专家组意见表

总体判断拟开设专业（方向）是否可行	√是 □否
理由： 1. 增设信息与计算科学专业（数据科学方向）的必要性。在当前大数据时代，利用前沿的数据科学和机器学习方法，探索自然科学、工程、经济、社会等多学科的交叉领域已成必然趋势。科学计算与机器学习的融合，已经成为当前应用数学发展的最重要前沿领域。计算数学中的很多基础算法，例如数值代数、最优化方法、微分方程数值算法、Monte Carlo 方法等，已经为数据科学中机器学习算法所广泛采用。与此同时，数据科学相关计算方法在实际应用过程中，也推动了计算数学的发展。发展数据科学既更加适应未来的需要，又能充分利用“信息与科学计算系”的资源优势，尽快地把应用数学、计算数学和信息科学等学科中的有生力量调动起来以开展工作。 2. 增设信息与计算科学专业（数据科学方向）符合就业和人才市场的需求。数据经济的高速发展，导致大数据人才的培养数量和速度远远达不到产业规模的增速，专业人员奇缺。北京大学在数据科学的教学方面强调基础和应用的结合，鼓励学生在数学模型、算法和计算机实现等方面全面发展。学生应该具备扎实的理论基础，同时也具备扎实的实践能力。以有效地避免目前这种数学专业学生缺乏实践能力，计算机专业的学生缺乏理论功底的尴尬局面。这样的人才将会得到信息产业的广泛欢迎。 3. 人才培养目标定位准确，培养方案科学合理。该方向致力于培养掌握数学、计算机、统计等数据科学相关领域基础理论知识，以及数据建模、机器学习、并行与分布式计算、统计推断等方法和技术，从事数据建模、数据分析与挖掘算法等问题的研究和大数据系统开发的研究型和技术型人才。课程设置围绕上述培养目标，强调毕业生应扎实地掌握数学、统计、信息科学的基础知识和数据建模、数据分析与挖掘等基本技能，具备独立学习的能力、初步的研究能力以及较强的适应不同社会职业需要的能力。 4. 师资力量和办学条件较好。北京大学学科种类齐全，综合优势明显，共有国家一级重点学科 18 个，国家二级重点学科 25 个。在大数据交叉学科建设方面的优势学科众多，如：数学、统计学、计算机、软件；应用经济学、社会学、新闻传播学、心理学；基础医学、临床医学、口腔医学、公共卫生与预防医学等。校企合作基础：与百度、腾讯、联想、方正、中移动、联通等知名企业开展了大数据相关的研究和人才培养合作，这些企业可为科学研究和人才培养提供真实的“大数据”和大数据运行环境。以上都为数据科学方向的设立奠定了基础。信息与计算科学专业（数据科学方向）由北京大学数学学院信息与计算科学系从事数据研究的优秀教师参与课程设计与教学，其中包括 1 名中国科学院院士、及多位长江学者、国家杰出青年基金获得者。自 2015 年起，北京大学数学科学学院已在国内高校之中率先设立数据科学与大数据技术本科专业。很多基础课程已经有很好的资源，新兴专业课程也已经具备相当的研究基础。 增设信息与计算科学专业（数据科学方向）可以满足我国对该方向人才的需求，开设该方向有较强的必要性和可行性，其培养目标明确、培养方案科学合理，拥有较好的专业教学师资力量和教学；实验设备、场地，学校支持力度高。建议加快批准该方向的开设和招生。	

拟招生人数与人才需求预测是否匹配		√是 □否												
本专业（方向）开设的基本条件是否符合教学质量国家标准	教师队伍	√是 □否												
	时间条件	√是 □否												
	经费保障	√是 □否												
专家签字：														
<table border="1"> <thead> <tr> <th>姓名</th> <th>所在单位</th> <th>签名</th> </tr> </thead> <tbody> <tr> <td>陈大岳</td> <td>北京大学</td> <td></td> </tr> <tr> <td>鄂维南</td> <td>北京大学</td> <td></td> </tr> <tr> <td>张平文</td> <td>武汉大学</td> <td></td> </tr> </tbody> </table>			姓名	所在单位	签名	陈大岳	北京大学		鄂维南	北京大学		张平文	武汉大学	
姓名	所在单位	签名												
陈大岳	北京大学													
鄂维南	北京大学													
张平文	武汉大学													

9. 医学类、公安类专业相关部门意见

(应出具省级卫生部门、公安部门对增设专业意见的公函并加盖公章)