

# 北京大学信息科学技术学院

## 数据科学与大数据技术专业培养方案

### 一、专业简介

数据科学与大数据技术是研究与数据的表示、处理（计算）、系统和智能有关的基础学科，是人类社会信息化、智能化发展的重要基础。本学科是在现代数学和物理、信息论、计算机软硬件等技术发展的基础上形成的多学科交叉的宽口径专业。传统的研究领域包括数据科学、数据管理、大数据技术与系统、人工智能等。进入 21 世纪，随着互联网、云计算、大数据、物联网和人工智能的飞速发展，更加拓宽了数据科学与大数据技术的研究范围，为该学科的发展注入了新的活力。

数据科学与大数据技术由信息科学技术学院和计算机学院共同建设。本专业的突出特色是“重视数理基础、强化系统能力、鼓励交叉创新”，坚持以学生为本，坚持通识和专业教育想融合的培养模式，培养数据科学与大数据领域基础理论扎实、在数据科学与大数据前沿开拓创新能力强、同时具有国际视野的未来领军人才。

### 二、培养目标

培养学生成为具有“引领未来、守正创新”的精神，具有国际视野和爱国敬业意识，具有“基础厚实、理工交叉、乐于探究、勇于创新”特点，能够成为新一代大数据系统与技术、人工智能、计算机软件和理论、计算机系统结构、计算机应用等领域引领数据科学与大数据技术发展创新的领军人才。

通过通识与专业相结合的教育，使学生具备坚实的数学、计算机、数据科学、人工智能等计算机软硬件基础知识，系统地掌握数据科学和大数据技术的理论和方法，受到良好的科学思维与科学实践研究的训练，具有探索、发现、分析和解决问题的能力，以及知识自我更新和不断创新的能力，为引领数据科学与大数据技术发展奠定基础。培养的学生具有正确的人生观和价值观，具有良好的人文和科学素养，具有独立思考、阅读、写作、表达等能力和国际化视野。

### 三、培养要求

本专业本科毕业生可在科研机构、高等院校、企业事业单位从事数据科学与大数据技术领域的研究、教学、开发、管理工作；也可继续攻读数据科学与大数据技术、计算机科学与技术、软件工程、智能科学与技术和其它相关学科的研究生学位。具体要求包括以下几个方面：

**【1】专业基础：**掌握数据科学与大数据技术领域所需要的数学、计算机、智能等专业基础

知识，具有较强的文献阅读、写作和外语交流能力，能够综合应用上述能力解决科学研究和/或实际工程开发问题。

【2】问题研究：能够基于科学原理，采用科学方法，运用系统思维和创新思维，针对实际工程科学应用和未来产业发展，提出新问题、新方法和新系统，体现创新能力。

【3】问题分析：能够应用数学、物理、计算机、网络等基本原理，分析未知问题的可能解决方案，结合文献研究、原理探索和独立思考，给出创新性的解决方案。

【4】解决问题：能够结合专业培养所获得的综合设计和实践能力，对解决方案的原理进行理论评估、实际测试和原理验证，并有能力开发出解决方案的原型系统，在实际环境中开展验证和演示。

【5】社会责任：能够在应用科学研究和实际工程开发中，自觉关注科学、技术和工程对人类社会可持续发展的影响，包括对环境、健康、安全、法律、伦理以及文化的影响，自觉遵守职业道德和规范，并履行应承担的责任。

【6】团队合作：具有较强的组织能力、沟通能力、表达能力和人际交往能力，能够在团队协作中发挥积极的作用，具有承担项目管理和团队负责的主动精神和能力。

【7】终身学习：具有自主学习和终身学习的意识和能力，具有较强的面向未知问题的主动探索精神和能力。

## 四、毕业要求及授予学位类型

本专业学生在学期间，须修满培养方案规定的 148 学分，方能毕业。达到学位要求者授予理学学士学位。

具体毕业要求包括：

1、公共基础课程：51 学分	1-1 公共必修课：39 学分
	1-2 通识教育课：12 学分
2、专业必修课程：59 学分	2-1 专业基础课：19 学分
	2-2 专业核心课：34 学分
	2-3 毕业论文（设计）：6 学分
3、选修课程：38 学分	3-1 专业选修课：18 学分
	3-2 自主选修课：20 学分

## 五、课程设置

1. 公共基础课程：45-51 学分

1-1 公共必修课：33-39 学分

课号	课程名称	学分	周学时	实践总学时	选课学期
03835xxx	大学英语	2-8			按大学英语教研室要求选课
	思想政治理论必修课	19			按马克思主义学院要求选课

	劳动教育课			32	按学校要求选课
04830041	计算概论 A	3	4	32	一上
04830050	数据结构与算法 A	3	4	32	二上
60730020	军事理论	2	2	0	一上
——	体育系列课程	4			全年，按体育教研室要求选课

可替代课程列表：

课号	课程名称	学分	周学时	实践总学时	替代课程
04830530	计算概论 A（实验班）	3	4	32	计算概论 A
04830540	数据结构与算法 A（实验班）	3	4	32	数据结构与算法 A

注：相关课程均可以由同名的实验班课程进行替代（下同）。

## 1-2 通识教育课：12 学分

通识教育课程有四个系列（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和通选课两部分课程，修读总学分为 12 学分。具体要求如下：

- （1）至少修读一门“通识教育核心课程”，且在四个课程系列中每个系列至少修读 2 学分；
- （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：59 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
00132511	高等数学 A（I）	5	6	32	一上
00132512	高等数学 A（II）	5	6	32	一下
00132611	线性代数 A（I）	4	5	32	一上
00132612	线性代数 A（II）	4	5	32	一下
04830010	信息科学技术概论	1	2	0	一上

可替代课程列表：

课号	课程名称	学分	周学时	实践总学时	替代课程
00132301	数学分析（I）	5	6	32	高等数学 A（I）
00132302	数学分析（II）	5	6	32	高等数学 A（II）
00132321	高等代数（I）	5	6	32	线性代数 A（I）
00132323	高等代数（II）	4	5	32	线性代数 A（II）

## 2-2 专业核心课：34 学分

课号	课程名称	学分	周学时	实践总学时	选课学期
04831750	程序设计实习	3	4	32	一下
新开课	人工智能基础	3	3	16	一下
新开课	离散数学基础	3	3	0	二上
04833040	计算机系统导论	5	4	0	二上
04832363	计算机系统导论讨论班	0	2	32	二上
04833050	算法设计与分析	5	4	0	二下
04832580	算法设计与分析（研讨型小班）	0	2	32	二下
22530003	AI 中的数学	3	3	0	二下
04834630	Python 程序设计与数据科学导论	3	3	0	二下
00131480	概率统计 A	3	3	0	三上
04834920	计算机视觉导论	3	3	0	三下
新开课	自然语言处理	3	3	16	三下

可替代课程列表：

课号	课程名称	学分	周学时	实践总学时	替代课程
04833400	离散数学与结构（I）	3	4	0	离散数学基础
04834040	人工智能引论	3	3	16	人工智能基础

## 2-3 毕业论文：6 学分

## 3. 选修课程：38 学分

### 3-1 专业选修课：18 学分

课号	课程名称	学分	周学时	实践总学时	选课学期
00132513	高等数学 A（III）	4	5	32	二上
04831730	机器学习概论	3	3	8	三下
04830230	计算机图形学	3	3	0	二下
04830320	数字图像处理	3	3	0	三下
04834760	智能机器人系统导论	3	3	0	三上
04834520	强化学习	3	3	0	三上
04834770	数值分析	3	3	0	二上
04833410	凸分析与优化方法	3	3	0	三下
00135460	数理统计	3	3	0	三上
04630790	数据科学导引	3	3	0	三上
00136720	大数据分析中的算法	3	3	0	三上
04834260	操作系统	4	5	32	三上/下
04830220	数据库概论	3	3	0	三下

04834210	计算机网络	4	5	32	三上/下
04832240	并行与分布式计算导论	3	3	0	三下
04830410	信息安全引论	2	2	0	四上

可替代课程列表:

课号	课程名称	学分	周学时	实践总学时	替代课程
00132304	数学分析 (III)	4	5	32	高等数学 A (III)
04833420	机器学习	3	3	0	机器学习概论
00130280	计算方法 B	3	3	0	数值分析

### 3-2 自主选修课: 20 学分 (全校课程均可)

课号	课程名称	学分	周学时	实践总学时	选课学期
04834200	编译原理	4	5	32	三上/下
04834220	软件工程	4	5	32	二下/三上
04830080	代数结构与组合数学	3	3	0	二下
04830070	集合论与图论	3	3	0	三上
04830090	数理逻辑	3	3	0	三上
04830260	理论计算机科学基础	3	3	0	三下
04833900	密码学基础	3	3	0	三上
04834230	软件测试导论	3	3	0	三上
04833020	软件分析技术	3	3	0	三上
04830140	计算机组织与体系结构	3	3	0	三上/下
04830145	计算机组织与体系结构实习	2	2	32	三上/下
04830100	数字逻辑设计	3	3	2	二下/三上
04832520	并行程序设计原理	2	2	8	二下
04834100	信息科学中的数学	2	2	0	二上
04830030	科技交流与写作	2	2	8	二下
04830340	JAVA 程序设计	2	2	2	二下
04834940	量子计算	3	3	0	三上
04834980	量子信息	3	3	0	三下
04831210	信息论	2	2	0	三下
04831800	数字媒体技术基础	2	2	4	三下
04830290	面向对象技术引论	2	2	0	三下
04830760	数字信号处理 (含上机)	3	4	16	三下
04830270	程序设计语言概论	2	2	4	四上
04830310	人机交互	2	2	0	四上
04830510	语言统计分析	2	2	0	四上
04831880	初等数论及其应用	3	3	0	四上
04831890	现代信息检索导论	2	2	0	四上

## 六、其他

### 1. 荣誉学位要求

针对愿意充分发展个人兴趣、积极开阔国际视野,追求更高科学和工程学位或学习体验的同学,本专业提供了实验班和荣誉课程系列(Honor Track),完成此系列课程中规定课程的学习,并达到以下相应要求的学生,将可以申请荣誉学士学位,评定通过后,学生将获得学校统一颁发的荣誉证书。

【1】思想品德好,在校期间没有受过任何纪律处分。

【2】已获得所修专业的学士学位授予资格。

【3】前7个学期总平均绩点位于本专业本科毕业生的前30%。

【4】前7个学期,完成以下不少于六门荣誉课程的学习要求,且成绩达到优秀( $\geq 85$ 分)。

课号	课程名称	学分	周学时	实践总学时	选课学期
04830530	计算概论 A (实验班)	3	4	32	一上
04831760	程序设计实习 (实验班)	3	4	32	一下
04830540	数据结构与算法 A (实验班)	3	4	32	二上
04833400	离散数学与结构 (I)	3	4	0	二上
04833440	计算理论导论	3	3	0	二下
04833420	机器学习	3	3	0	二下
04833060	算法设计与分析 (实验班)	5	6	32	二下
04834580	软件工程 (实验班)	4	5	32	二下/三上
04834490	操作系统 (实验班)	4	5	32	三上/下
新申请	编译原理 (实验班)	4	5	32	三上/下
04834480	计算机网络 (实验班)	4	5	32	三上/下
04830221	数据库基础 (实验班)	3	3	0	三下

【5】申请学生应当参与本科生科学研究项目或申请获得“研究课程”学分,且结题答辩成绩 $\geq 85$ 分。

【6】毕业论文评价等级优秀。

### 2. 港澳台学生和留学生学分与选课要求

【1】港澳台学生和留学生除免修课程外,学分完成要求均与本科生要求一致。

【2】免修课程的替代要求如下:

免修全校公共必修课程中的思想政治理论以及军事理论课,需选修“与中国有关的课程”<sup>21</sup>学分替代。

### 3. 特色培养项目或奖励要求

鼓励对研究感兴趣的、未来打算攻读博士的同学,申请加入计算机学科教育部“拔尖计划 2.0”项目,在学院组织的各项科研活动之外,还可以申请国际合作与交流、暑期科研实

习、参加学术会议等经费支持。

#### **4. 优秀毕业生（Honor Student Award）奖励要求**

- 【1】** 思想品德好，在校期间没有受过任何纪律处分。
- 【2】** 已获得所修专业的学士学位授予资格。
- 【3】** 前 7 个学期总平均绩点位于本专业本科毕业生的前 30%。
- 【4】** 前 7 个学期，以下四门课程中至少选修三门，且每门优秀（ $\geq 85$  分）：操作系统、编译原理、计算机组织与体系结构、计算机网络。
- 【5】** 毕业论文优秀。

注：优秀毕业生证书由信息科学技术学院与专业共建单位计算机学院共同发放。

#### **5. 其他课程方面规定**

- 【1】** 相同课名或者授课内容相近的课程为互斥课程，不能重复计算学分。外院系选修的同名课程也不能计算学分。如果有疑问，请提前和教务老师确认。
- 【2】** 大学英语所修学分不足 8 学分(或免修)的同学需通过专业或通识选修课程补齐学分。

七、数据科学与大数据专业课程地图





