



《大数据行业的兴起》

|  |  |
| --- | --- |
| **专 业** | **数据科学与大数据技术** |
| **姓 名** | **董宇臻** |
| **日 期** | **2022/9/7** |

**摘 要**

以教育部2015-2019年公布本科专业备案与审批结果数据为基础，通过调查研究与统计分析方法，对数据科学与大数据技术专业发展的状况与趋势进行研究。主要研究包括四个方面：对数据科学与大数据技术专业设立的时代背景进行综述研究；对我国数据科学与大数据技术专业的办学规模进行统计分析；对数据科学的特点进行总结，并且对发展趋势进行预测；对数据科学与大数据技术专业人才培养特点进行分析。厘清数据科学与大数据技术的历史发展脉络。

**目 录**

[**第1章 专业研究内容介绍** - 1 -](#_Toc113465161)

[1.1 数据科学与大数据技术专业是什么 - 1 -](#_Toc113465162)

[1.2 数据科学与大数据技术专业有哪些特点 - 1 -](#_Toc113465163)

[**第2章 专业主要课程内容** - 2 -](#_Toc113465164)

[2.1 数据科学与大数据技术专业学什么 - 2 -](#_Toc113465165)

[**第3章 专业学习建议与指导** - 3 -](#_Toc113465166)

[3.1 数据科学与大数据技术专业学习指导 - 3 -](#_Toc113465167)

[**第4章 目前专业的前沿研究方向** - 4 -](#_Toc113465168)

[4.1 数据科学的特点与发展趋势 - 4 -](#_Toc113465169)

[**第5章 专业本科生毕业去向** - 5 -](#_Toc113465170)

[5.1 关于深造 - 5 -](#_Toc113465171)

[5.2 关于就业 - 5 -](#_Toc113465172)

[**结 论** - 6 -](#_Toc113465173)

[**参考文献** - 7 -](#_Toc113465174)

**第1章 专业研究内容介绍**

## 1.1 数据科学与大数据技术专业是什么

大数据是一个你可能不了解但一定听说过的热门词汇，大数据并非是指数据很大，而更侧重于海量数据。大数据的价值并不在于这些数据本身，更注重的是通过对数据进行采集、存储、分析、处理等环节挖掘出来的信息。随着大数据的火爆，与大数据相关的专业也应运而生，数据科学与大数据技术专业就是其中之一。

该专业是一门以统计学、数学和计算机为三大支撑的交叉学科，涉及人工智能、数据可视化、机器学习等领域的知识，不同的学校在课程设置上会有不同的侧重点。但总体来说，数据科学与大数据技术专业的设置是为了培养具有大数据思维、能够熟练掌握大数据分析和大数据应用技术的高级大数据人才，既重视学习面向大数据应用的数学、统计学、计算机科学与计算等学科基础知识，又强调训练面向大数据挖掘与分析的技能，注重培养从数据管理、系统开发和海量数据分析与挖掘这三个大数据应用的主要层面出发，对实际问题进行分析和解决的能力。

## 1.2 数据科学与大数据技术专业有哪些特点

数据科学与大数据技术专业最典型的特点就是交叉领域较广，需要学生掌握数学、统计学和计算机三大学科的知识。数据分析对统计学知识要求较高，数据挖掘要求数学能力较强。该专业部分课程内容稍有难度，整体来说，大部分学生能够入门，但要想把大数据学精深，还需要自身不懈努力。

数据科学与大数据技术专业另一个特点是应用范围广泛，各行各业都需要大数据人才，比如金融、军事、互联网等行业。这一特点决定了大数据知识只是一个工具，要想解决实际问题，还需要掌握相关行业的专业知识。

**第2章 专业主要课程内容**

## 2.1 数据科学与大数据技术专业学什么

具体包括高等代数、随机过程、数学分析、解析几何、常微分方程、离散数学、多元统计分析、数理分析、大数据概论、智能计算、机器学习与数据挖掘、人工智能、神经网络与深度学习、hadoop大数据开发、机器学习、数据机构、Python与大数据分析、分布式与并行计算、数据库原理及应用、Java程序设计、R语言、大数据存储与管理、Matlab应用等。

课程设置分为专业基础课、专业核心课和专业选修课三大类。不同的学校对大数据人才培养的侧重点不同，在课程设置上会根据本校人才培养方案稍有变动。数据科学与大数据技术专业注重对学生计算思维能力的培养，强调从数理逻辑到不确定性思维的培养，要求具备较强的大数据处理的系统构建、算法模型、应用推理等认知和实践能力。该专业总体发展方向分为三大类：数据分析、数据平台建设和数据治理。数据分析方向要求掌握数据采集、准备、存储、计算和处理方面的知识，掌握SQL、R语言、数据可视化工具、常见的算法模型及统计方法，对逻辑思维和表达能力要求较高。平台建设方向主要掌握软件工程和项目管理知识。数据治理方向需要熟悉各种数据治理的框架。

**第3章 专业学习建议与指导**

## 3.1 数据科学与大数据技术专业学习指导

整体来说，数据科学与大数据专业基础知识较容易，但部分核心课程难度较大，要想学好，需要学生付出较多的时间和精力，对学生素质要求较高。因为该专业是新兴专业，难免存在部分学校课程设置不合理的情况，所以学生在选择学校时需提前了解学校的人才培养方案，多做对比。大数据行业非常注重实践能力和行业经验，所以选择本专业的学生要利用寒暑假的时间多参与实习实践，尽量选择专业对口岗位，为以后就业积攒经验。

**第4章 目前专业的前沿研究方向**

## 4.1 数据科学的特点与发展趋势

传统的数据理论与技术主要来源于近代统计学与其他学科的交叉融合，例如，数理统计学、社会统计学、计量经济学、统计物理学、生物统计学、教育统计学、信息统计学、卫生统计学等。传统统计学是研究数据的采集、建模、分析、可视化与表达理论与技术的自然科学分支。传统的数据主要是“小数据”，传统的统计学研究的对象也主要限制在“小数据”集合上，由于“小数据”依赖人工采集使得掌握总体的成本过高，所以传统统计学发展强调用样本去估计总体的思维方式，样本的采集抽样、分布假设、估计理论、推断理论等均建立在数学的公理化体系下，这保证了传统统计学的严谨性与合理性。

直到“大数据”的基础设施和有关技术的出现，传统数据思维和方法体系开始调整。在计算机算力受限制的情况下，传统统计学的算法效率大部分无法跟上快速增长的数据规模变化。在此种情况下，一些以效率优先的算法逐渐进入到应用层面，某些算法舍去一定数学上的严谨性，甚至内部机制与原理并不完全清楚，但运算简洁清晰，能够使用现有算力对海量数据进行有效处理，并且以结果为导向，通过算法形式和参数的迭代，在预测或对问题分类的准确性上不断优化，使这类专门处理“大数据”的方法突破学科理论研究的局限，迅速部署到商业和工业领域，并经过若干年的发展，开始反过来影响科学研究的学科分类，以及高等教育的新专业形成。

数据科学被提出，集合并统一了传统的统计学与跨学科的大数据技术，重新界定了自然科学的一个分支领域，数据科学的历史使命，不但需要继续发扬统计学已有的研究理论与交叉领域，而且要在新的数据规模和方法领域发展出系统的理论体系，并将提取和解释新旧问题的科学本质与运行机制，建立新的方法论体系去指导和帮助信息技术条件下其他学科的发展和跨越。数据科学将成为新的“横截面”学科，有机会与其他所有自然科学和社会科学各学科交叉，只要其他学科有被信息化的条件和可能，就会产生大规模数据集的投影，进而就需要与数据科学相结合产生专门的问题集合与方法算法集合，新的交叉学科分支就会出现，数据科学的发展将会加速其他相关学科的信息化与数据化发展速度。

强大的应用层面需求也会急速推动数据科学的快速发展，使得未来20—30年出现数据科学发展的高速时期。目前，我国在这一领域和其他科技强国处在同一水平线和历史时期，同时因为我国在基础设施、制造业规模、政策导向、教育规模以及科技研发投入等方面具有明显的规模优势与先发态势，而且经过移动互联网、工业互联网和物联网的快速建设，中国已经拥有了世界上最大规模的网民数量、智能机器与终端设备数量，这已经产生了人类有史以来最大规模的数据储备规模，这一切都是中国加速建设成为数据强国的先决条件。在这种时代背景下，数据科学与大数据技术专业的建设、发展以及培养人才的大规模稳定输出，将有力保障我国现代化建设、产业升级、政府治理能力提升，以及制造业2025和民族复兴等伟大事业的成功。

**第5章 专业本科生毕业去向**

## 5.1 关于深造

数据科学与大数据技术专业旨在培养精通大数据分析与应用的高端人才，因此考研也是大数据专业学生比较好的选择。目前该专业学生深造方向分为两类：一类是继续从事大数据方面的研究，在大数据与产业结合领域寻找创新点，为自己开拓新的发展路线；第二类是选择人工智能方向，相对于大数据来说，人工智能有许多可以继续深入研究的领域，学生更容易找到自己的发展方向。大数据专业是以计算机学科为支撑的，所以在选择考研专业时也可以选择传统的计算机专业。总体来说，该专业考研可选择的范围较宽。另外，国外比国内的技术体系更为成熟，因此去国外学校深造，也是不错的选择。

## 5.2 关于就业

毕业生可以在政府部门、企事业单位和公司从事大数据分析、管理、研究和应用开发等方面的工作，具体分为管理类、分析类和研发类岗位。管理类岗位主要是产品经理和运营经理，负责大数据产品的规划与设计，结合业务需求，设计数据监控模型，搭建数据产品架构等内容；分析类包括分析类工程师和算法工程师，需要使用数据挖掘、机器学习等方法进行数据清洗和数据分析，为客户提供有价值的信息；研发类岗位主要分为架构工程师、开发工程师和运维工程师，负责基于Hadoop、spark等数据分析平台进行设计、开发、分布式计算等任务。该行业整体收入比较高，随着对人工智能以及大数据技术的人才需求逐步上升，未来可能会发展为就业前景最好的专业之一。

**结 论**

数据科学与大数据技术专业是顺应当今时代和历史潮流应运而生的高等院校新型本科专业，具有鲜明的新理论、新技术、新产业与新趋势特征，将成为我国夺取第四次工业革命制高点的关键人才输送阵地。在国家的有力统筹与部署下，数据科学将进入较长一段时间的快速发展期，将迎来人才、理论、技术与产业的质量和规模的集中爆发。在这一历史进程中，我国高等院校和相关行业中数据科学的快速发展，将进一步促进整个自然科学与社会科学的发展，激发我国整体科学研究水平与社会生产力发展的进步。

**参考文献**

[1] 孙书韬，朱立谷，李春芳.北美知名大学数据科学专业课程体系分析[J].中国教育信息化，2019(24).

[2] 王日芬，谢清楠，宋小康.国外数据科学研究的回顾与展望[J].图书情报工作，2016,60(14).

[3] 魏瑾瑞，蒋萍.数据科学的统计学内涵[J].统计研究，2014,31(5).