数据科学与大数据技术（信息技术）专业

本科人才培养方案

（专业代码080910T）

**一、培养目标**

本专业旨在培养具有正确的社会主义核心价值观以及规范使用语言文字的意识和应用能力，德、智、体、美全面发展，以数学与信息技术为基础，系统掌握数据科学与大数据技术等专业知识，具有搜集、整理、分析和呈现大数据的基本能力，能够熟练使用在大数据架构上的基本工具进行数据分析，为企业、咨询和研究机构解决实际问题的应用型大数据分析人才。

**二、毕业要求**

1. 掌握马克思主义、毛泽东思想与中国特色社会主义基本理论，具有良好的思想品德、个人修养以及健康的心理和体魄，具有良好的人文社会科学素养、职业道德和心理素质，社会责任感强。

2. 掌握统计专业所必须的数学基础知识，理解如何使用数学语言表达统计思想。

3. 具有扎实的数据科学理论功底，系统掌握数据科学和大数据专业知识，了解国内外国际前沿动态。

4. 具有较强的计算机及互联网应用能力，掌握高效的处理大数据的方法，掌握大数据建模和分析的基础理论，能够从大数据中获取有价值的知识。

5. 具备解决交叉学科实际问题的能力。具备使用数据科学思维方法、信息技术和大数据技术对经济、金融、管理等领域数据信息进行收集和分析处理，完成相关领域的实际工作。

6. 具有较好的外语应用能力，能阅读本专业的外文材料，具有宽广的国际视野和跨文化交流与合作能力。

7. 信息职业素质的培养。熟悉大数据专业领域相关政策及法律、法规，能够在本专业领域实践活动中理解并遵守职业道德和职业规范。

8. 掌握体育运动的一般知识和基本方法，具有健康的体魄、良好的心理素质、稳定的情绪和坚强的意志，形成良好的体育锻炼习惯。

**三、培养特色**

1．实现统计与信息类学科的交叉培养。本专业（方向）的培养注重多学科的协同与创新，由统计学院和信息学院共同合作完成，充分体现交叉学科的特点优势。注重学科间的相互合作与团队协作，不仅加强各种统计软件在大数据中的应用，而且注重计算机软件在大数据架构下的应用。

2．注重实践能力的培养与业界的结合。本专业（方向）在重视统计理论与方法应用教学的同时，更强调实践与社会需求相结合，以国内知名企业作为实践平台，重视学生的实践活动与各行业的海量数据分析相结合。

3．本专业（方向）与国内国外有关学校深入广泛合作，与港澳台和国外知名大学有充分的交流与合作，有着突出的教育国际化的特点。

**四、核心课程**

数学分析、高等代数、数据科学的概率基础、面向对象程序设计（JAVA）、数据科学的计算机基础、经济学原理、数据科学导论、数据科学的统计基础、数据库原理与应用、抽样技术与应用、回归分析、分布式计算、时间序列分析、数据结构、统计编程基础、程序设计语言（Python）、数据仓库与数据挖掘、机器学习、大数据预处理、数据科学算法

**五、学制和学位**

本专业基本学制为四年，实行弹性修业年限制度，学生在校修业年限可以提前至三年或延长至六年，修满规定的学分准予毕业。符合学士学位授予条件者，授予理学学士学位。

**六、学分一览表**

1. **总学分一览表**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **总学分** | **教学方式** | | | | **理论与实践教学比例（%）** | |
| 173 | 课堂教学环节 | 148 | 理论教学 | 115 | 理论教学 | 66 |
| 实验教学 | 33 | 实践教学 | 34 |
| 课外教学环节 | 25 | 实习、军训等 | 25 |

1. **课堂教学学时分配表**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **课程类别** | **门数** | **总学时** | **总学分** | **比例（%）** |
| 通识教育必修课 | 19 | 992 | 58 | 39 |
| 学科基础课 | 9 | 416 | 26 | 18 |
| 专业必修课 | 7 | 352 | 22 | 15 |
| 专业选修课 |  |  | 22 | 15 |
| 通识教育选修课 |  |  | 14 | 9 |
| 个性化课程 |  |  | 6 | 4 |
| **合计** |  |  | **148** | **100** |

1. **实践教学环节一览表**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **类别** | **实践环节** | **学期安排** | **周数** | **学时** | **学分** |
| 实习类 | 军训（含军事理论课） | 1 | 3 |  | 4 |
| 认知实习 | 4 | 3 |  | 2 |
| 专业实习 | 6 | 3 |  | 2 |
| 毕业实习 | 8 | 8 |  | 4 |
| 毕业论文（设计） | 8 | 8 |  | 4 |
| 素质类 | 创新学分 |  |  |  | 2 |
| 第二课堂 |  |  |  | 2 |
| 思想政治类 | 思想道德修养与法律基础 | 2 |  |  | 1 |
| 毛泽东思想与中国特色社会主义理论体系概论 | 1 |  |  | 1 |
| 形势与政策 | 4 |  |  | 1 |
| 中国近代史纲要 | 3 |  |  | 1 |
| 马克思主义基本原理概论 | 4 |  |  | 1 |
| 计算机类 | 面向对象程序设计（JAVA） | 1 | 16 | 32 | 2 |
| 专业实验类 | 数学分析Ⅰ | 1 | 16 | 16 | 1 |
| 数学分析Ⅱ | 2 | 16 | 16 | 1 |
| 数据科学的概率基础 | 3 | 16 | 16 | 1 |
| 数据科学的计算机基础 | 2 | 16 | 32 | 2 |
| 数据科学的统计基础 | 4 | 16 | 16 | 1 |
| 数据库原理与应用 | 4 | 16 | 32 | 2 |
| 回归分析 | 5 | 16 | 16 | 1 |
| 分布式计算 | 5 | 16 | 16 | 1 |
| 时间序列分析 | 6 | 16 | 16 | 1 |
| 数据结构 | 3 | 16 | 16 | 1 |
| 统计编程基础 | 3 | 16 | 32 | 2 |
| 数据仓库与数据挖掘 | 5 | 16 | 16 | 1 |
| 机器学习 | 6 | 16 | 16 | 1 |
| 大数据预处理 | 6 | 16 | 32 | 2 |
| 操作系统 | 5 | 16 | 16 | 1 |
| 计算机网络技术与应用 | 3 | 16 | 16 | 1 |
| 数据采集与存储 | 4 | 16 | 16 | 1 |
| 数据科学算法 | 7 | 16 | 16 | 1 |
| 优化方法 | 7 | 16 | 16 | 1 |
| 统计预测与决策（双语） | 7 | 16 | 16 | 1 |
| 非结构数据分析与建模 | 6 | 16 | 16 | 1 |
| 图像分析技术及应用 | 5 | 16 | 16 | 1 |
| 分析型大数据系统 | 5 | 16 | 16 | 1 |
| 程序设计语言（Python） | 4 | 16 | 16 | 1 |
| 数据可视化 | 6 | 16 | 32 | 2 |
| 自然语言处理 | 6 | 16 | 16 | 1 |
| **合计** | |  |  |  | **58** |

**七、本科学分制指导性教学计划表**

**八、专业经典阅读书目及期刊目录**

（1）专著

[1] 范明，范宏建.数据挖掘导论 [M]. 北京：人民邮电出版社，2010.

[2] [Robert Tibshirani](https://book.douban.com/search/Robert%20Tibshirani/)，[Trevor Hastie](https://book.douban.com/search/Trevor%20Hastie/)，[Jerome Friedman](https://book.douban.com/search/Jerome%20Friedman/). 统计学习基础 [M]. 北京: 电子工业出版社, 2004

[3] 李航. 统计学习方法 [M]. 北京:清华大学出版社，2012.

[4] David Barber. Bayesian Reasoning and Machine Learning [M].英国:剑桥大学出版社，2011.

[5] [陈涛](https://www.amazon.cn/s/ref=dp_byline_sr_book_4?ie=UTF8&field-author=%E9%99%88%E6%B6%9B&search-alias=books), [张旭](https://www.amazon.cn/s/ref=dp_byline_sr_book_5?ie=UTF8&field-author=%E5%BC%A0%E6%97%AD&search-alias=books),[崔杨](https://www.amazon.cn/s/ref=dp_byline_sr_book_6?ie=UTF8&field-author=%E5%B4%94%E6%9D%A8&search-alias=books),[刘海平](https://www.amazon.cn/s/ref=dp_byline_sr_book_7?ie=UTF8&field-author=%E5%88%98%E6%B5%B7%E5%B9%B3&search-alias=books). Python自然语言处理 [M]. 北京:人民邮电出版社, 2014.

[6] 周志华，机器学习[M]. 北京:清华大学出版社.2016.

[7] Bishop. Pattern Recognition and Machine Learning [M]. 美国:Springer, 2007.

[8] 汤森.蔡, 金融时间序列分析 [M].北京：人民邮电出版社.2012.

[9] 陈宝林，最优化理论与算法[M].北京:清华大学出版社.2005

[10] Daphne Koller, Nir Friedman, Probabilistic Graphic Models [M].美国:MIT 出版社，2009.

（2）论文集

[1] Conference on Neural Information Processing Systems [C]. 美国:MIT press， 每年

[2] International Conference on Machine Learning [C]. 美国: ACM，每年

[3] IEEE Conference on Computer Vision and Pattern Recognition [C]. 美国：IEEE, 每年

[4] Conference on Empirical Methods in Natural Language Processing [C]. 美国: ACL, 每年

[5] Conference on Knowledge Discovery and Data Mining [C]. 美国: ACM,每年

（3）期刊

[1] 统计学、数学、计算机科学，Journal of Machine Learning Research, MIT press.

[2] 统计学、数学、计算机科学，IEEE Transactions on Knowledge and Data Engineering， IEEE.

[3] 统计学、数学、计算机科学，IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE.

[4] 统计学、数学，Journal of the American Statistical Association, Taylor & Francis

[5] 统计学、数学，Biometrika, Oxford University Press.

[6] 统计学、数学，Annals of Statistics,IMS.

1. **九、培养目标与毕业要求矩阵图**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **培养目标**  **毕业要求** | **具有正确的社会主义核心价值观以及规范使用语言文字的意识和应用能力** | **扎实的数学、数据科学、大数据专业理论知识** | **运用大数据技术进行数据分析的能力** | **具有较强的解决实际问题的能力** |
| 思想政治素质和人文科学素养 | H | M | L | L |
| 数学基础知识 | H | H | H | M |
| 数据科学和大数据专业知识 | H | H | H | H |
| 计算机应用及高效的大数据处理能力 | H | H | H | H |
| 交叉学科综合应用与创新能力 | H | M | H | H |
| 外语应用能力 | H | M | H | M |
| 信息职业素质 | H | M | L | M |
| 身心素质 | H | H | H | H |

注： H表示高相关性，M表示中相关性，L表示低相关性。

1. **十、毕业要求与课程体系矩阵图**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **毕业要求**  **课程名称** | 掌握马克思主义、毛泽东思想和中国特色社会主义理论体系 | 掌握通识教育类、学科基础类、专业基础类、专业类知识及相关学科知识 | 能将所学知识用于解释计算机专业领域及相关领域的现象和问题 | 能够恰当使用数学工具、现代编程工具和方法 | 掌握一门外语，具备较好的语言表达和沟通协调能力 | 具有团队协作意识 | 具有自主学习和终身学习意识 | 具有人文素养、科学精神和社会责任感 |
| 毛泽东思想和中国特色社会主义理论体系概论 | **√** |  |  |  |  |  |  |  |
| 思想道德修养与法律基础 | **√** |  |  |  |  |  | **√** | **√** |
| 大学生心理健康 | **√** |  |  |  |  |  | **√** | **√** |
| 中国近现代史纲要 | **√** |  |  |  |  |  |  |  |
| 马克思主义基本原理概论 | **√** |  |  |  |  |  |  |  |
| 形势与政策 | **√** |  |  |  |  |  | **√** |  |
| 大学综合英语 | **√** |  |  |  | **√** |  |  |  |
| 数学分析 | **√** | **√** |  |  |  |  |  |  |
| 高等代数 | **√** | **√** |  |  |  |  |  |  |
| 数据科学的概率基础 | **√** | **√** |  |  |  |  |  |  |
| 体育 | **√** |  |  |  |  |  |  | **√** |
| 面向对象程序设计（JAVA） | **√** |  | **√** | **√** |  |  |  |  |
| 应用写作 | **√** |  |  |  | **√** |  |  | **√** |
| 数据科学的计算机基础 | **√** |  | **√** | **√** |  | **√** |  |  |
| 经济学原理 | **√** |  |  |  |  | **√** |  |  |
| 数据科学导论 | **√** |  | **√** | **√** |  |  |  |  |
| 抽样技术与应用 | **√** |  | **√** |  |  |  |  |  |
| 数据科学的统计基础 | **√** | **√** | **√** |  |  |  |  |  |
| 数据库原理与应用 | **√** |  | **√** | **√** |  |  |  |  |
| 回归分析 | **√** |  | **√** |  |  |  |  |  |
| 分布式计算 | **√** |  | **√** |  |  |  |  |  |
| 时间序列分析 | **√** |  | **√** |  |  |  |  |  |
| 数据结构 | **√** |  | **√** | **√** |  |  |  |  |
| 统计编程基础 | **√** |  | **√** | **√** |  |  |  |  |
| 数据仓库与数据挖掘 | **√** |  |  | **√** |  |  |  |  |
| 数据科学算法 | **√** |  | **√** | **√** |  |  |  |  |
| 机器学习 | **√** |  | **√** | **√** |  |  |  |  |
| 大数据预处理 | **√** |  | **√** | **√** |  |  |  |  |
| 操作系统 | **√** |  | **√** | **√** |  |  |  |  |
| 计算机网络技术与应用 | **√** |  | **√** | **√** |  |  |  |  |
| 数据采集与存储 | **√** |  | **√** | **√** |  |  |  |  |
| 分布式计算 | **√** |  | **√** | **√** |  |  |  |  |
| 优化方法 | **√** | **√** | **√** |  |  |  |  |  |
| 统计预测与决策  (双语） | **√** |  | **√** |  |  |  |  |  |
| 图像分析技术及应用 | **√** |  | **√** | **√** |  |  |  |  |
| 非结构数据分析与建模 | **√** |  | **√** |  |  |  |  |  |
| 分析型大数据系统 | **√** |  | **√** | **√** |  |  |  |  |
| 安全通论 | **√** |  | **√** | **√** |  |  |  |  |
| 程序设计语言（Python） | **√** |  | **√** | **√** |  |  |  |  |
| 数据可视化 | **√** |  | **√** |  |  |  |  |  |
| 自然语言处理 | **√** |  | **√** |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

1. **十一、课程修读及培养流程图**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1. **第一学年** | | **第二学年** | | **第三学年** | | **第四学年** | |
| **一** | **二** | **三** | 四 | 五 | **六** | 七 | 八 |

体育课II

体育课III

体育IV

体育I

马克思主义基本原理概论

通识教育必修课

优化方法

统计预测与决策

安全通论

毕业实习及毕业论文（设计）

面向对象程序设计（JAVA）

大学综合英语II

大学综合英语I

大学综合英语III

通识教育

学科基础课课

专业必修课

专业教育

程序设计语言

（Python）

计算机网络技术与应用

数据采集与存储

数据科学算法

机器学习、

大数据预处理

数据仓库与数据挖掘

数据结构、

统计编程基础

时间序列分析

回归分析

分布式计算

数据科学导论

抽样技术与应用

数据科学的统计基础  
数据库原理与应用

经济学原理

数据科学的计算机基础

中国近现代史纲要

高等代数、数据科学的概率基础

数学分析II

数学分析I

毛泽东思想和中国特色社会主义理论体系概论

第二课堂、创新创业教育；通识教育选修课；个性化选修

非结构数据分析与建模

自然语言处理

数据可视化

图像分析技术及应用

操作系统

分析型大数据系统

专业选修课

思想道德修养与法律基础、大学生心理健康