



微信扫码刷题



免费约直播领资料



免费订阅考试提醒

数据分析师进阶必备 6 大数学利器

一、函数、变量、方程、图

学什么

从最简单的学起，比如线性方程、二项式定理及其性质。

对数、指数、多项式函数、有理数

基本几何定理、三角恒等式

实数和复数及其基本性质

级数、求和与不等式

绘图、制图、笛卡尔系和极坐标系、圆锥曲线

怎么用

要理解百万量级数据排序后再搜索会快多少，首先要理解什么是二分法搜索，要理解二分法搜索，就要懂得什么是对数，什么是递归方程式。还有，如果想分析时间序列，就要了解周期函数和指数衰减的概念。

在哪儿学

数据科学的数学 - Coursera

代数简介 - edX

可汗学院 - 代数

二、统计学

学什么

统计学可是数据分析师的必备技能。想做数据分析必须要有坚实的统计学与概率论基础，这点就不用多说了。除了新出现的神经网络机器学习，传统的机器学习其实就是统计学习，比如李航的统计学习方法讲的就是机器学习原理。统计学的内容非常广泛，我们只要关注最核心的概念就可以了。

数据摘要与描述性统计：集中趋势、方差、协方差、相关性

概率论基础：基本概念、期望、概率演算、贝叶斯定理、条件概率



扫二维码下载 环球网校移动课堂 APP

移动学习 职达未来



微信扫码刷题



免费约直播领资料



免费订阅考试提醒

概率分布函数: 均匀分布、正态分布、二项分布、卡方分布、学生 t 分布、中心极限定理

采样、测量、误差、随机数生成

假设检验: A/B 检验、置信区间、P 值

方差分析(ANOVA)、t 检验

线性回归、正则化

怎么用

面试的时候就会用得上, 相信我, 作为准数据科学家, 如果把统计学的概念搞得清清楚楚、明明白白, 一定能让面试官刮目相看。当上了数据科学家, 统计学更是常用的工具。

在哪儿学

用 R 学统计学专业—Coursera, 杜克大学

Python 数据科学 - 统计学与概率论—edX, 加利福尼亚大学

商务统计与分析专业课—Coursera, 莱斯大学

三、线性代数

学什么

Facebook 上的朋友推荐、Spotify 上的歌曲推荐, 通过深度学习把自拍照片转换为萨尔瓦多·达利风格的人像画, 试问这些操作的共同点是什么? 其实就是它们都离不开矩阵和矩阵代数的知识。

线性代数这一重要的数学分支, 研究的是机器学习算法如何从数据流中获取有价值信息。下列是必学的线性代数知识:

矩阵和向量的基本性质: 标量乘法、线性变换、转置、共轭、秩与行列式

内积与外积、矩阵乘法规则及其算法、逆矩阵

特殊矩阵: 方阵、单位矩阵、三角矩阵、稀疏矩阵、密集矩阵、单位向量、对称矩阵、厄米矩阵(又称自共轭矩阵)、斜厄米矩阵和酉矩阵

矩阵分解概念: LU 分解、高斯消元法、求解 $Ax = b$ 线性方程组

向量空间、基向量、扩张空间、正交性、线性最小二乘法



扫二维码下载 环球网校移动课堂 APP

移动学习 职达未来



微信扫码刷题



免费约直播领资料



免费订阅考试提醒

特征值、特征向量、对角化与奇异值分解(SVD)

怎么用

使用降维技术实现主成分分析时，要用奇异值分解来缩减维度，让数据集参数变得更少。所有神经网络算法都使用线性代数处理网络结构和学习操作。

在哪儿学

线性代数：从基础到精通 —edX, UT Austin

机器学习中的数学知识：线性代数 —Coursera, 帝国理工学院, 伦敦

四、微积分

学什么

在上大学的时候，微积分是最让人头疼的课程，不过，在数据科学和机器学习领域里，微积分可是无处不在，最简单的普通最小二乘法问题的分析解决方案离不开微积分，神经网络中用于学习新模式的反向传播里也少不了微积分。可以说，微积分是你的技能库里最有价(zhi)值(qian)的技能。

下列是要掌握的微积分知识点：

单变量函数、极限与连续、可微性

均值定理、不定式和洛必达法则

最大值与最小值

乘积和链式法则

泰勒级数、无限级数求和与积分

积分中值定理与基本公式，定积分与不定积分方程式

Beta 和 Gamma 函数

多变量函数、极限与连续、偏导数

常微分方程与偏微分方程的基础知识(不必了解过于高级的内容)

怎么用



扫二维码下载 环球网校移动课堂 APP

移动学习 职达未来



微信扫码刷题



免费约直播领资料



免费订阅考试提醒

理解逻辑回归算法需要微积分的知识, 比如怎样通过“梯度下降”找到最小损失函数。要了解梯度下降的机制, 就会用到微积分的概念 - 梯度、导数、极限和链式法则。

在哪儿学

大学预备课 - 微积分 — edX, TU Delft

可汗学院 - 微积分

机器学习中的数学知识: 多变量微积分 — Coursera, 帝国理工学院, 伦敦

五、离散数学

学什么

探讨数据科学的数学知识时, 很少会谈及离散数学这个话题, 但几乎所有现代数据科学都需要计算系统的支持, 而这些系统的核心恰恰是离散数学。大一学生在学习离散数学时会被告知: 初学者一定要掌握日常分析项目使用的算法和数据结构。离散数学核心知识点如下:

集合、子集、幂集

计数、组合、可数性

基本证明技巧: 归纳法、反证法

归纳、演绎和命题逻辑基础

数据结构基础: 堆栈、队列、图、数组、哈希表、树

图形性质: 连通分支、度、最大流与最小割、染色图

递归关系与方程

函数增长率与大 O 符号法

怎么用

对于任何社交网络分析, 都需要理解图形性质与快速算法来查找和遍历网络。不管选择哪种算法都要理解算法的时空复杂性, 比如随着输入数据大小的变化, 运算所需的时空需求如何变化, 这时通常会用到大 O 符号法。

在哪儿学



扫二维码下载 **环球网校移动课堂 APP**

移动学习 职达未来



微信扫码刷题



免费约直播领资料



免费订阅考试提醒

计算机专业课: 离散数学简介—Coursera, Univ. of California San Diego

数学思想简介—Coursera, 斯坦福大学

精通离散数学: 集合与数学逻辑—Udemy

六、最优化、运筹学

学什么

这一部分主题说得都是与应用数学相关的知识, 最常用的是计算机科学原理、控制论、运筹学等。理解这些概念对机器学习实践来说非常重要。

实际上, 每种机器学习算法都要在限制条件下实现误差估计最小化, 这就是优化。要学习的内容如下:

最优化基础: 如何规划命题

最大值、最小值、凸函数、全局解

线性规划、单纯形算法

整数规划

约束规划、背包问题

怎么用

用最小二乘法损失函数解决简单线性回归问题一般能得到较为精确的分析解, 但在解决逻辑回归问题时就不行了。要想理解其中的原因, 就要理解最优化里的凸性概念。这个概念还告诉我们, 在绝大多数机器学习问题里要能够接受近似解, 这是个毋庸置疑的事实。

在哪儿学

商业分析最优化—edX, 麻省理工学院

离散优化—Coursera, 墨尔本大学

确定性优化—edX, 佐治亚理工学院

结语

即使数学不好也不用过于担心, 更不用迷茫无助。想成为资深数据分析师要学的东西很多, 如果平时不怎么应用数学知识, 就更要下些工夫。但是, 如今这个时代最好的地方就是网上有超多优秀的资源, 比如各类视频教程。只要花些时间, 投入精力, 就可以找到适合自己的学习资源。



扫二维码下载 环球网校移动课堂 APP

移动学习 职达未来

环球网校

www.hqwx.com
咨询热线: 400-678-3456



微信扫码刷题



免费约直播领资料



免费订阅考试提醒

我可以保证，即便在大学学过这些内容，现在重温或学习新的数学知识以后，你会发现这些时间与精力没有白费，一定能一点一点地开始理解数据分析与机器学习项目背后所隐藏的旋律。数据分析师进阶必备 6 大数学利器，这就是进阶成为数据科学家所要迈出的一大步。

环球网校
www.hqwx.com



扫二维码下载 **环球网校移动课堂 APP**

移动学习 职达未来