

序号	课程名称	课程内容	补充材料	注释
1	课程介绍	1、课程介绍 2、大数据分析基础 3、数据科学的基本内容 4、数据科学、机器学习和人工智能	1、数据嗨客的《数据科学导引》课程第1章	
2	大数据分析的python基础（上）	1、python 基础内容回顾 2、numpy 简介 3、scipy 简介	1、数据嗨客的《打数据分析的Python基础》勘测 2、《利用python进行数据分析（第二版）》第2-4章 3、《Python数据科学手册》第1-2章	
3	大数据分析的python基础（下）	1、pandas 简介 2、seaborn 和 matplotlib 简介 3、sklearn 简介	1、《利用python进行数据分析》第5章 2、《Python数据科学手册》第3-5章 3、数据嗨客的《Pandas数据分析》课程	
4	回归模型	1、线性回归与非线性回归 2、梯度下降法（GD） 3、正则化 4、随机梯度下降法（SGD）	1、数据嗨客的《数据科学导引》课程第3章	
5	梯度下降法的优化	1、适应性学习率 2、AdaGrad、AdaDelta、RMSProp、Adam、AMSGrad、……		
6	特征工程与模型评估	1、特征选择 2、模型评估	1、周志华《机器学习》第2章 2、数据嗨客的《数据科学导引》课程第7-8章	
7	数据预处理	1、数据标准化 2、特征编码 3、特征离散化 4、缺失值处理	1、数据嗨客的《数据科学导引》课程第2章 2、《python机器学习》第4章	
8	逻辑回归	1、逻辑回归 2、梯度下降	2、数据嗨客的《数据科学导引》课程第4.3章	

9	决策树	1、基本流程 2、划分选择 3、剪树	1. 周志华《机器学习》 4.1—4.3 章 2、数据嗨客的《数据科学导引》课程第 4.5 章	
10	神经网络上	1、神经元模型 2、感知机与多层网络	1、周志华《机器学习》5.1-5.2 章	
11	神经网络下	1、反向传播算法（BP 神经网络） 2、全局最小与局部极小 3、进阶	1、周志华《机器学习》5.3-5.6 章	
12	支持向量机	1、间隔与支持向量 2、拉格朗日乘子法 3、核函数与核方法	1、周志华《机器学习》第 6 章 2、数据嗨客的《数据科学导引》课程第 4.1 章	
13	贝叶斯分类器	1、贝叶斯决策论 2、极大似然估计 3、朴素贝叶斯分类器	1、周志华《机器学习》7.1-7.3 章 2、数据嗨客的《数据科学导引》课程第 4.4 章	
14	集成模型	1、Boosting 2、AdaBoost 算法 3、随机森林算法	1. 周志华《机器学习》 8.1-8.3 章 2、数据嗨客的《数据科学导引》课程第 5 章	
15	聚类模型	1、聚类任务 2、K 均值算法（k-means 算法）	1. 周志华《机器学习》 9.1-9.4 章 2、数据嗨客的《数据科学导引》课程第 7-8 章	
16	降维	1、k 近邻算法 2、主成分分析 3、核化线性降维	1. 周志华《机器学习》 10.1-10.4 章 2、数据嗨客的《数据科学导引》课程第 4.2、9 章	
17	实战工具	1、XGBoost 2、lightGBM		

参考网站

数据嗨客：<http://hackdata.cn/#>

参考书籍

周志华《机器学习》
《利用 python 进行数据分析》
《Python 数据科学手册》
《Python 机器学习》