

机器学习工程师

1 开班：课程学习指南

第一章 课程学习指南

1.1 【课程】课程学习指南

2 导论：机器学习入门与算法总览

第一章 机器学习介绍与算法一览

1.1 【课程】本章概述

1.2 【课程】机器学习是什么？举几个栗子告诉你！

1.3 【课程】机器学习的基本概念和典型应用场景

1.4 【课程】机器学习的2个环节和4个阶段

1.5 【课程】机器学习的评估方法与评估指标

1.6 【课程】机器学习算法一览与可视化

1.7 【课程】如何学好《机器学习工程师微专业》

1.8 【案例】案例分享：学完能有多厉害？机器学习经典算法的Scikit-Learn实现

3 算法与案例：线性回归与逻辑回归

第一章 线性回归 (Linear Regression) 与逻辑回归 (Logistic Regression)

1.1 【课程】本章概述

1.2 【课程】线性模型

1.3 【课程】线性回归(Linear Regression)算法详解与经典的房价预测案例

1.4 【课程】广义线性模型

1.5 【课程】逻辑回归 (Logistic Regression, LR)算法详解

1.6 【课程】工程应用经验分享，回答你最关心的实操问题

1.7 【案例】手把手带你使用Python实现线性回归算法

1.8 【案例】手把手带你使用Python实现逻辑回归算法

4 算法与案例：树模型

第一章 决策树 (Decision Tree) 与分类问题

1.1 【课程】决策树模型概述与简史

1.2 【课程】决策树算法基本流程与最佳属性选择方法

1.3 【课程】剪枝与控制过拟合

1.4 【案例】手把手带你实现基于决策树的收入分类与可视化

第二章 回归树与随机森林 (Random Forest)

2.1 【课程】本章概述

2.2 【课程】连续值和缺省值的处理

2.3 【课程】回归树(Regression Tree)模型详解、

2.4 【课程】Bagging思想与随机森林(Random Forest, RF)模型详解

2.5 【案例】使用Iris经典数据集练习决策树与随机森林的建模流程和可视化操作

2.6 【案例】使用房价预测经典案例练习回归树建模和随机森林建模流程

5 算法与案例：支持向量机

第一章 线性支持向量机 (Support Vector Machine)

1.1 【课程】本章概述

1.2 【课程】二分类线性可分支持向量机

1.3 【课程】二分类线性不可分支持向量机

1.4 【课程】多分类支持向量机

1.5 【课程】SVM工具包：LibSVM、SVMLight、Scikit-Learn

1.6 【案例】Kaggle案例分享：基于SVM分类的信用卡欺诈检测

第二章 非线性支持向量机 (Support Vector Machine)

2.1 【课程】本章概述

2.2 【课程】SVM约束优化问题与SVM对偶形式

2.3 【课程】核函数(Kernel Function)与核技巧

2.4 【课程】非线性支持向量机与SOM(Sequential Minimal Optimization)优化算法详解

2.5 【课程】支持向量回归(Support Vector Regression)算法详解

2.6 【案例】Kaggle案例分享：基于SVM回归的房价预测

6 工程与实战：特征工程与模型调优

第一章 特征工程的原理详解

1.1 【课程】本章概述

1.2 【课程】特征工程的意义与工业界的实际应用

1.3 【课程】基本数据处理：采集、清洗、采样

1.4 【课程】4种常见数据类型的特征工程操作要点

1.5 【课程】3种特征选择方法

第二章 特征工程的实战演示

2.1 【课程】本章概述

2.2 【案例】使用Titanic经典案例练习基本数据处理操作

2.3 【案例】手把手教你10种常见的特征工程操作

2.4 【案例】手把手教你3类常见的特征选择操作

第三章 模型调优与模型融合

3.1 【课程】本章概述

3.2 【课程】数据预处理与模型的选择和调参

3.3 【课程】模型状态与模型调优

3.4 【课程】模型融合的重要性与5类常见操作方式

3.5 【案例】使用Pima Indians Diabetes经典案例完成4类模型融合操作演练

7 算法与案例：无监督算法

第一章 聚类算法

1.1 【课程】本章概述

1.2 【课程】聚类问题及典型应用场景介绍

1.3 【课程】K均值 (K-means) 算法和 K中心点 (K-medoids)算法详解

- 1.4 【课程】K-means算法的扩展：高斯混合模型 (Gaussian mixture model, GMM) 和Soft K-means
- 1.5 【课程】层次聚类 (Hierarchical Clustering)算法详解

第二章 降维算法

- 2.1 【课程】本章概述
- 2.2 【课程】降维问题及典型应用场景介绍
- 2.3 【课程】线性降维算法：主成分分析 (Principal Component Analysis, PCA)算法详解
- 2.4 【课程】非线性降维算法：等距特征映射 (Isometric Feature Mapping, IsoMap)算法详解
- 2.5 【课程】多维标度 (Multidimensional Scaling, MDS)算法详解

第三章 关联规则挖掘

- 3.1 【课程】本章概述
- 3.2 【课程】关联规则挖掘问题及典型应用场景介绍
- 3.3 【课程】购物篮分析和频繁集挖掘
- 3.4 【课程】关联数据挖掘最经典算法：频繁项集算法 (Apriori Algorithm)详解
- 3.5 【课程】提升效率：频繁模式生长 (Frequent Pattern-Growth, FP-Growth)算法详解

8 算法与案例：概率机器学习

第一章 最大熵与 EM 算法

- 1.1 【课程】本章概述
- 1.2 【课程】统计学基础回顾：先验概率与后验概率
- 1.3 【课程】统计学基础回顾：极大似然估计(MLE)
- 1.4 【课程】信息与熵
- 1.5 【课程】最大熵原理与模型
- 1.6 【课程】最大期望 (Expectation Maximization, EM) 算法详解
- 1.7 【案例】手把手带你使用Python实现EM算法

第二章 贝叶斯网络 (Bayesian Networks)

- 2.1 【课程】本章概述
- 2.2 【课程】朴素贝叶斯 (Naive Bayes, NB) 算法详解
- 2.3 【课程】贝叶斯网络与有向分离
- 2.4 【课程】马尔科夫模型 (Markov Model, MM) 详解与实际应用场景介绍
- 2.5 【案例】手把手带你使用Python实现朴素贝叶斯算法

第三章 隐马尔科夫模型 (Hidden Markov Model)

- 3.1 【课程】本章概述
- 3.2 【课程】隐马尔科夫模型 (Hidden Markov Model, HMM) 详解
- 3.3 【课程】使用隐马尔科夫链解决第1类问题：概率计算问题 (Evaluation)
- 3.4 【课程】使用隐马尔科夫链解决第3类问题：状态预测问题 (Recognition)
- 3.5 【课程】使用隐马尔科夫链解决第3类问题：参数学习问题 (Training)
- 3.6 【课程】隐马尔科夫链的应用：词性标注及其算法详解
- 3.7 【案例】手把手带你使用Python实现基于HMM的词性标注

第四章 主题模型 (Topic Model)

- 4.1 【课程】本章概述
- 4.2 【课程】主题模型与LDA (Latent Dirichlet Allocation, LDA) 算法原理
- 4.3 【课程】主题模型4类算法详解：Unigram Model、Mixture of Unigram Model、LDA、pLSA
- 4.4 【案例】使用LDA模型破解希拉里“邮件门”的秘密

9 算法与案例：推荐系统

第一章 推荐系统

- 1.1 【课程】本章概述
- 1.2 【课程】推荐系统的广泛应用与评估指标
- 1.3 【课程】基于内容的推荐 (User-Based Recommendation) 系统算法详解
- 1.4 【课程】基于物品的协同过滤推荐 (Item-Based Collaborative Filtering Recommendation) 系统算法详解
- 1.5 【课程】基于用户的协同过滤 (User-Based Collaborative Filtering Recommendation) 系统算法详解
- 1.6 【课程】协同过滤优缺点与冷启动问题的解决方案
- 1.7 【课程】隐语义模型 (Latent Factor Model, LFM) 算法详解
- 1.8 【案例】手把手带你完成协同过滤算法的隐语义模型 (LFM) 实现
- 1.9 【案例】手把手带你使用Scikit-Surprise库搭建电影推荐系统

10 工程与实战：机器学习工具库

第一章 机器学习常用工具库实战：sklearn

- 1.1 【课程】本章概述
- 1.2 【课程】最好用的机器学习库：Scikit-Learn 使用指南
- 1.3 【课程】Scikit-Learn 解决机器学习问题的一般路径
- 1.4 【案例】天池大航杯“制造扬中”电力AI大赛：使用线性模型完成AI电力能耗预测
- 1.5 【案例】天池大航杯“制造扬中”电力AI大赛：使用树模型完成AI电力能耗预测
- 1.6 【案例】手把手带你使用朴素贝叶斯算法搭建语种检测模型
- 1.7 【案例】手把手带你使用机器学习算法完成中文文本分类

第二章 机器学习高级工具库实战：XGBoost与LightGBM

- 2.1 【课程】本章概述
- 2.2 【课程】集成模型家族与工具库
- 2.3 【课程】XGBoost 参数介绍与通用代码模板
- 2.4 【课程】LightGBM 参数介绍与通用代码模板
- 2.5 【案例】手把手带你使用XGBoost用法速查表完成7个练习
- 2.6 【案例】手把手带你学习sklearn与XGBoost配合使用的4大要点
- 2.7 【案例】手把手带你使用LightGBM用法速查表完成5个练习
- 2.8 【案例】手把手带你学习sklearn与LightGBM配合使用的3大要点
- 2.9 【案例】Kaggle便利店销量预测经典案例

第三章 机器学习实战案例深度解析

- 3.1 【课程】本章概述
- 3.2 【案例】Kaggle房价预测比赛优胜案例代码详解
- 3.3 【案例】Data Castle大学生助学金精准资助预测比赛代码详解
- 3.4 【案例】Data Castle大学生助学金精准资助预测比赛金牌团队模型深度解析

11 数据科学比赛：TOP1方案深度解析

第一章 机器学习实战案例深度解析

- 1.1 【课程】本章概述
- 1.2 【案例】Kaggle房价预测比赛优胜案例代码详解
- 1.3 【案例】Data Castle大学生助学金精准资助预测比赛代码详解
- 1.4 【案例】Data Castle大学生助学金精准资助预测比赛金牌团队模型深度解析