机器学习

第1课 绪论

1. 课程介绍：

2. 机器学习

2.1 概念

2.2 学科定位

2.3 定义

2.4 学习

针对经验和一系列的任务和一定的表现的衡量，如果随之经验的积累，针对定义好的任务可以挺高表现，就说计算机具有学习能力。

3. 机器学习的应用

4 .应用的例子

基本概念

1. 基本概念：训练集，测试集，特征值，监督学习，非监督学习，半监督学习，分类，回归

2. 概念学习：

定义：概念学习是指从某个布尔函数的输入输出训练样例中推断出该布尔函数

3. 享受运动的例子

4. 重要的概念

训练集，训练样例，用来进行训练，也就是产生模型或者算法的数据集

测试集，测试样例，用来专门进行测试已经学习好的模型或算法的数据集

特征向量：属性的集合，通常用一个向量来表示，附属于一个实例的标志

标记

正例

反例

5. 研究美国硅谷房价

6. 分类：目标标记为类别数据

回归：目标标记为连续性数值

7.例子 研究肿瘤的良性，

有监督学习：训练集有类别

无监督学习：

半监督学习：

8. 机器学习步骤框架

8.1 把数据拆分为了训练集和测试集

8.2 用训练集和训练集的特征向量进行训练算法

8.3 用学习来的算法运用在测试集上评估算法（可能要涉及到调整参数，用到验证集）

第2讲 决策树

3.1 熵的概念

第3讲 决策树的应用

1. Python

2. Python的机器学习的库：scikit-learn

3.使用scikit-learn

3.1 安装 pip安装

3.2 anaconda

4.1 最近邻规则分类（k-Nearest Neighbor）KNN 算法

4.1 综述

4.2 例子

4.3 算法详述

为了判断未知实例

4.4 算法的优缺点