

<<机器学习导论>>

图书基本信息

书名：<<机器学习导论>>

13位ISBN编号：9787111265245

10位ISBN编号：7111265246

出版时间：2009-6

出版时间：机械工业出版社

作者：Ethen Alpaydin

页数：272

译者：范明, 咎红英, 牛常勇

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

前言

自从有计算机以来，人们就希望计算机能够学习。

然而，机器学习真正取得实质性进展，能够成功地解决一些实际问题，并最终成为一个学科分支还是近20余年的事。

对于许多问题，我们的前人和先行者已经知道如何求解。

例如，欧几里德告诉我们可以用辗转相除法求两个整数的最大公约数；Dijkstra告诉我们如何有效地求两点之间的最短路径；Hoare向我们展示了怎样将杂乱无章的对象快速排序.....对于这些问题，我们清楚地知道求解步骤。

因此，让计算机求解这些问题只需要设计算法和数据结构、进行编程，而不需要让计算机学习。

还有一些事情，人们可以轻而易举地做好，但是却无法解释清楚我们是如何做的。

例如，尽管桌子千差万别、用途各异，但是我们一眼就能看出某个物体是否是桌子；尽管不同的人的手写阿拉伯数字大小不一、笔画粗细不同，但是我们还是可以轻易识别一个数字是不是；尽管声音时大时小、有时可能还有点沙哑，但是我们还是可以不用费力气地听出熟人的声音。

诸如此类的例子不胜枚举。

对于这些问题，我们不知道求解步骤。

因此，让计算机来做这些事就需要让计算机学习。

我们知道桌子不是木材和各种材料的随机堆砌，手写数字不是像素的随机分布，熟人的声音也不是各种声波的随机混合。

现实世界总是有规律的。

机器学习正是从已知实例中自动发现规律，建立对未知实例的预测模型；根据经验不断提高，不断改进预测性能。

这是关于机器学习这一主题全面论述的教科书，适合作为高等院校计算机相关专业高年级本科生和研究生机器学习入门课程的教材。

该书涵盖了监督学习、贝叶斯决策理论、参数方法、多元方法、维度归约、聚类、非参数方法、决策树、线性判别式、多层感知器、局部模型：隐马尔可夫模型、分类算法评估和比较、组合多学习器以及增强学习。

作者对来自统计学、模式识别、神经网络、人工智能、信号处理、控制和数据挖掘等不同领域的机器学习问题和学习方法进行了统一论述。

<<机器学习导论>>

内容概要

机器学习的目标是对计算机编程，以便使用样本数据或以往的经验来解决给定的问题。已经有许多机器学习的成功应用，包括分析以往销售数据来预测客户行为，人脸识别或语音识别，优化机器人行为以便使用最少的资源来完成任务，以及从生物信息数据中提取知识的各种系统。为了对机器学习问题和解进行统一的论述，本书讨论了机器学习在统计学、模式识别、神经网络、人工智能。

信号处理、控制和数据挖掘等不同领域的应用。

对所有学习算法都进行了解释，以便读者可以容易地将书中的公式转变为计算机程序。

本书可作为高等院校计算机相关专业高年级本科生和研究生的教材，也可供研究机器学习方法的技术人员参考。

本书对机器学习的定义和应用实例进行了介绍，涵盖了监督学习。

贝叶斯决策理论。

参数方法、多元方法、维度归约、聚类、非参数方法、决策树。

线性判别式、多层感知器，局部模型、隐马尔可夫模型。

分类算法评估和比较，组合多学习器以及增强学习等。

作者简介

Ethem Alpaydin，是土耳其伊斯坦布尔博阿齐奇大学计算机工程系的教授。于1990年在洛桑联邦理工学院获博士学位，之后先后在美国麻省理工和伯克利大学工作和进行博士后研究。

Ethem博士主要从事机器学习方面的研究，是剑桥大学的《The Computer Journal》杂志编委和Elsevier

<<机器学习导论>>

书籍目录

出版者的话 中文版序 译者序 前言 致谢 符号表 第1章 绪论 1.1 什么是机器学习 1.2 机器学习的应用实例 1.2.1 学习关联性 1.2.2 分类 1.2.3 回归 1.2.4 非监督学习 1.2.5 增强学习 1.3 注释 1.4 相关资源 1.5 习题 1.6 参考文献 第2章 监督学习 2.1 由实例学习类 2.2 VC维 2.3 概率逼近正确学习 2.4 噪声 2.5 学习多类 2.6 回归 2.7 模型选择与泛化 2.8 监督机器学习算法的维 2.9 注释 2.10 习题 2.11 参考文献 第3章 贝叶斯决策定理 3.1 引言 3.2 分类 3.3 损失与风险 3.4 判别式函数 3.5 效用理论 3.6 信息值 3.7 贝叶斯网络 3.8 影响图 3.9 关联规则 3.10 注释 3.11 习题 3.12 参考文献 第4章 参数方法 4.1 引言 4.2 最大似然估计 4.2.1 伯努利密度 4.2.2 多项密度 4.2.3 高斯(正态)密度 4.3 评价估计: 偏倚和方差 4.4 贝叶斯估计 4.5 参数分类 4.6 回归 4.7 调整模型的复杂度: 偏倚/方差两难选择 4.8 模型选择过程 4.9 注释 4.10 习题 4.11 参考文献 第5章 多元方法 5.1 多元数据 5.2 参数估计 5.3 缺失值估计 5.4 多元正态分布 5.5 多元分类..... 第6章 维度灾难 第7章 聚类 第8章 非参数方法 第9章 决策树 第10章 线性判别式 第11章 多层感知器 第12章 局部模型 第13章 隐马尔可夫模型 第14章 分类算法评估和比较 第15章 组合多学习器 第16章 增强学习

章节摘录

第1章 绪论 1.1 什么是机器学习 随着计算机技术的发展,我们现在已经拥有存储和处理海量数据以及通过计算机网络从远程站点访问数据的能力。目前大多数的数据存取设备都是数字设备,记录的数据也很可靠。以一家连锁超市为例,它拥有遍布全国各地的数百家分店,并且在为数百万顾客提供数千种商品的零售服务。销售点的终端设备记录每笔交易的详细资料,包括日期、顾客识别码、购买商品和数量、消费总额等。

这是典型的每日几个一字节的数据。只有分析这些数据,并且将它转换为可以利用的信息时,这些存储的数据才能变得有用,例如做预测。

我们不能确切地知道哪些人比较倾向于购买哪种特定的商品,也不知道应该向喜欢读海明威作品的人推荐哪位作者。如果我们知道,我们就不需要任何数据分析;我们只管供货并记录下编码就可以了。但是,正因为 we 不知道,所以才只能收集数据,并期望从数据中提取这些问题或相似问题的答案。我们确信存在某种过程,可以解释我们所观测到的数据。尽管我们不清楚数据产生过程(例如顾客行为)的细节,但是,我们知道数据产生不是完全随机的。人们并不是去超市随机购买商品。当人们买啤酒时,也会买薯片;夏天买冰淇淋,而冬天则为Glthwein买香料。数据中存在确定的模式。

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>