

# 《机器学习项目开发实战》

## 书籍信息

版次：1

页数：

字数：

印刷时间：2016年07月01日

开本：128开

纸张：胶版纸

包装：平装

是否套装：否

国际标准书号ISBN：9787115429513

## 编辑推荐

本书向读者展示了，如何利用简单的算法和技术，从数据中学习，构建更聪明的.NET应用，以解决现实世界中更广泛的问题。读者将在熟悉的Visual Studio环境中编码各个项目，使用.NET环境中适合于机器学习的F#语言开发机器学习项目。如果你对F#还很陌生，本书将教会你入门所需的知识。如果你已经熟悉了F#，本书将是你在机器学习领域实践该语言的新的机会。

在一系列令人着迷的项目中，读者将学到：

- \* 从头开始构建一个光学字符识别（OCR）系统
- \* 编写一个通过例子学习的垃圾邮件过滤器
- \* 使用F#强大的类型提供程序与外部资源接口（在本书中是来自R语言的数据分析工具）
- \* 将数据转换为信息量更大的特征，并用它们作出精准的预测
- \* 在不知道目标的情况下找出数据中的模式
- \* 用回归模型预测数值

[显示全部信息](#)

## 内容简介

本书通过一系列有趣的实例，由浅入深地介绍了机器学习这一炙手可热的新领域，并且详细介绍了适合机器学习开发的Microsoft F#语言和函数式编程，引领读者深入了解机器学习的基本概念、核心思想和常用算法。书中的例子既通俗易懂，同时又十分实用，可以作为许多开发问题的起点。通过对本书的阅读，读者无须接触枯燥的数学知识，便可快速上手，为日后的开发工作打下坚实的基础。本书适合对机器学习感兴趣的.NET开发人员阅读，也适合其他读者作为机器学习的入门参考书。

## 作者简介

Mathias Brandewinder是Microsoft F#

有价值专家（MVP），住在加州旧金山，在那里他为Clear Lines Consulting工作。作为一名当之无愧的数学极客，他很早就对构建模型帮助其他人利用数据做出更好的决策感兴趣。他拥有商业、经济 and 运营研究等多个硕士学位，在到达硅谷之后不久便爱上了编程。从.NET刚出现时开始，他就专业开发软件，为各行各业开发业务应用程序，重点是预测模型和风险分析程序。

目录

目录

第1章 256级灰度 1

1.1 什么是机器学习 2

1.2 经典的机器学习问题：图像分类 3

1.2.1 挑战：构建一个数字识别程序 3

1.2.2 机器学习中的距离函数 5

目录

第1章 256级灰度 1

1.1 什么是机器学习 2

1.2 经典的机器学习问题：图像分类 3

1.2.1 挑战：构建一个数字识别程序 3

1.2.2 机器学习中的距离函数 5

1.2.3 从简单的方法入手 5

1.3 我们的第一个模型（C#版本） 6

1.3.1 数据集组织 6

1.3.2 读取数据 7

1.3.3 计算图像之间的距离 9

1.3.4 编写分类器 11

1.4 那么，如何知道程序有效？ 12

1.4.1 交叉验证 12

1.4.2 评估模型质量 13

1.4.3 改进模型 14

1.5 介绍用于机器学习的F# 15

- 1.5.1 使用F#交互执行进行实时脚本编写和数据研究 15
- 1.5.2 创建第一个F#脚本 18
- 1.5.3 剖析第一个F#脚本 19
- 1.5.4 创建函数管道 22
- 1.5.5 用元组和模式匹配操纵数据 23
- 1.5.6 训练和评估分类器函数 24
- 1.6 改进我们的模型 26
  - 1.6.1 试验距离的另一种定义 26
  - 1.6.2 重构距离函数 27
- 1.7 我们学到了什么 30
  - 1.7.1 在好的距离函数中找到什么 30
  - 1.7.2 模型不一定要很复杂 31
  - 1.7.3 为什么使用F#? 31
- 1.8 更进一步 32
- 第2章 垃圾邮件还是非垃圾邮件? 33
  - 2.1 挑战：构建一个垃圾邮件检测引擎 34
    - 2.1.1 了解我们的数据集 34
    - 2.1.2 使用可区分联合建立标签模型 35
    - 2.1.3 读取数据集 36
  - 2.2 根据一个单词决定 38
    - 2.2.1 以单词作为线索 38
    - 2.2.2 用一个数字表示我们的确定程度 39
    - 2.2.3 贝叶斯定理 40
    - 2.2.4 处理罕见的单词 42
  - 2.3 组合多个单词 42
    - 2.3.1 将文本分解为标记 42
    - 2.3.2 简单组合得分 43
    - 2.3.3 简化的文档得分 44
  - 2.4 实现分类器 45
    - 2.4.1 将代码提取到模块中 46
    - 2.4.2 文档评分与分类 47
    - 2.4.3 集合和序列简介 49
    - 2.4.4 从文档语料库中学习 51
  - 2.5 训练第一个分类器 53
    - 2.5.1

[显示全部信息](#)

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www.tushupdf.com](http://www.tushupdf.com)