

MLAPP 读书笔记 - 00 前言

A Chinese Notes of MLAPP, MLAPP 中文笔记项目

<https://zhuanlan.zhihu.com/python-kivy>

记笔记的人: [cycleuser](#)

2018年05月06日13:15:36

本文内容仅仅是读书笔记，并非对原文的忠实翻译，且不包含任何原文图片。

简介

电子形式的数据规模增长，带来了数据分析自动化方法的需求的增长。机器学习的目的是开发能从数据中自动识别模式的方法，然后用发现的模式去对未来数据进行预测等等。所以机器学习和统计与数据挖掘领域都很相关，但重点和术语方面有所不同。本书就是对这一领域进行详细介绍，包含了一些应用样例，这些样例的领域包括分子生物学、文本处理、计算机视觉和机器人。

目标读者

这本书的目标读者是高年级研究生或者计算机科学、统计学、电子工程、经济等其他有充足相关数学知识的本科生。要求读者熟悉多元积分、概率论、线性代数以及计算机编程。初步了解统计学会有帮助，不过不太熟悉也不用担心阅读本书会有困难。

本书并不仅仅是一个罗列不同启发式方法的菜谱式书籍，而是强调了机器学习中的基于模型为原则的角度。对于任意的给定模型，都可以用一系列不同算法。反过来说，任意的算法也都往往可以用于多种不同模型。这样实现了一种模块化，将模型和算法相互区分，对教学和工程来说都是好选择。

本书会经常用到图像化的语言来对模型进行简明直观的表达。除了有助于理解之外，图结构还有利于开发高效的算法。不过这本书的重点并不是图结构，而是一般意义上的概率建模。

本书中所提到的方法几乎都包含于一个叫做 PMTK 的 MATLAB 软件包里面，这个 PMTK 的意思就是概率建模工具箱（Probabilistic modeling toolkit）的英文缩写。PMTK 软件可以从 <https://github.com/probml/pmtk3> 下载，原来书中的链接 pmtk3.googlecode.com 提示失效了。pmtk 后面的这个3是指版本号，本书用的是 pmtk3，更多相关资源可以访问 <https://github.com/probml>，其中由代码、文档、图件等等。

关于 MATLAB 的介绍这里就不说了，建议大家使用开源的 MATLAB 替代品 Octave。

本书中的很多图片都是使用 PMTK 生成的，这些相关的代码也都在 PMTK 网站上可以找到。这部分内容大家自己在 PMTK 的网站上看吧。