Broadview®
www.broadview.com.cn

让“高高在上”的AI技术回归供人们学习和使用的层面

基于代码理解机器学习和深度学习在实际应用中的意义

程序员的AI书

从代码开始

张力柯　潘　晖　编著

中国工信出版集团

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

程序员的AI书

从代码开始

张力柯　潘　晖　编著

電子工業出版社
Publishing House of Electronics Industry
北京•BEIJING

内 容 简 介

随着 AI 技术的普及，如何快速理解、掌握并应用 AI 技术，成为绝大多数程序员亟需解决的问题。本书基于 Keras 框架并以代码实现为核心，详细解答程序员学习 AI 算法时的常见问题，对机器学习、深度神经网络等概念在实际项目中的应用建立清晰的逻辑体系。

本书分为上下两篇，上篇（第 1～4 章）可帮助读者理解并独立开发较简单的机器学习应用，下篇（第 5～9 章）则聚焦于 AI 技术的三大热点领域：推荐系统、自然语言处理（NLP）及图像处理。其中，第 1 章通过具体实例对 Keras 的机器学习实现进行快速介绍并给出整体概念；第 2 章从简单的神经元开始，以实际问题和代码实现为引导，逐步过渡到多层神经网络的具体实现上，从代码层面讲解神经网络的工作模式；第 3 章讲解 Keras 的核心概念和使用方法，帮助读者快速入门 Keras；第 4 章讲解机器学习中的常见概念、定义及算法；第 5 章介绍推荐系统的常见方案，包括协同过滤的不同实现及 Wide&Deep 模型等；第 6 章讲解循环神经网络（RNN）的原理及 Seq2Seq、Attention 等技术在自然语言处理中的应用；第 7～8 章针对图像处理的分类及目标识别进行深度讨论，从代码层面分析 Faster RCNN 及 YOLO v3 这两种典型识别算法；第 9 章针对 AI 模型的工程部署问题，引入 TensorFlow Serving 并进行介绍。

本书主要面向希望学习 AI 开发或者转型算法的程序员，也可以作为 Keras 教材，帮助读者学习 Keras 在不同领域的具体应用。

图书在版编目（CIP）数据

程序员的 AI 书：从代码开始 / 张力柯，潘晖编著. —北京：电子工业出版社，2020.2
ISBN 978-7-121-38270-3

Ⅰ. ①程… Ⅱ. ①张… ②潘… Ⅲ. ①人工智能－程序设计 Ⅳ. ①TP18

中国版本图书馆 CIP 数据核字（2020）第 014136 号

责任编辑：张国霞　　特约编辑：顾慧芳
印　　刷：三河市君旺印务有限公司
装　　订：三河市君旺印务有限公司
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　　邮编 100036
开　　本：787×980　1/16　印张：20　字数：430 千字
版　　次：2020 年 2 月第 1 版
印　　次：2020 年 2 月第 1 次印刷
印　　数：5000 册　　定价：109.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。

作者简介

张力柯

腾讯某AI实验室负责人、AI系统设计专家。在操作系统内核、网络安全、搜索引擎、推荐系统、大规模分布式系统、图像处理、数据分析等领域具有丰富的实践经验。

于美国德克萨斯大学圣安东尼奥分校获得计算机科学博士学位，曾先后在美国微软、BCG、Uber及硅谷其他创业公司担任研发工程师及项目负责人等。

潘　晖

阿里巴巴某算法中心小组负责人。在推荐系统、自然语言处理、图像处理、数据分析等领域具有丰富的实践经验。

于美国佛罗里达理工大学获得计算机科学博士学位，曾先后在中国微软、美团、腾讯从事算法研发和管理工作。发表过多篇论文，拥有多项专利，曾获得2018年腾讯互动娱乐事业群技术突破奖等奖项。

推荐序一

认识力柯兄多年，一直认为他是一员虎将——能用代码说话，便绝不打无谓的嘴仗；能用技术与产品直接证明，便绝不空谈“形式”和“主义”。这次，通过力柯兄写的这本书，一如既往地看到他“心有猛虎”的一面：直截了当、大开大合。

在机器学习或者说工业界 AI 火起来的这几年，程序员这一受众群体一直缺少优秀的教程。有些教程过于浅显，很难称其为“入门教程”，只能称其为科普书；有些教程则过于贴近理论推导，对夯实读者的数理基础大有裨益，也能给研究生提供参考，但对程序员来说，则理论有余而实践不足，常常令注重工程实践的他们一头雾水：毕竟不是每个程序员都有耐心、有必要一门一门地捡回微积分、概率统计、偏微分方程、线性代数和数值计算等基础学科的知识，再真正实现一个属于自己的模型。如何在数学理论和工程实践之间找到一个平衡点，让具有工程背景的广大读者从中获得实际的价值，而非进行简单的脑力或数学训练，这一直是我评价机器学习教程时最为看重的要素。现在，我有幸从力柯兄的成书中找到了这些要素，实乃幸事！

这是一本写给程序员看的机器学习指南。它有针对性地从程序员的视角切入（而非像市面上的大多数机器学习教程一样，从数学的角度切入），介绍了工业界流行的若干模型及应用场景，同时涵盖了神经网络的原理和基础实现、Keras 库的使用方法和 TensorFlow 的部署方案，可谓有的放矢。另外，本书章节不多，却简短有力。这不是一本科普读物，不存在浅尝辄止；也不是一本百科全书，不存在天书符号。这是一本有代码的书，是一本谈工程实现的书。我认为，这正是机器学习领域所缺少的那一类教程。

本书的上篇，让我不由得想起多年前力柯兄刚从硅谷回国高就时，与我围绕“怎样的面试题对于机器学习程序员是合适的”这一话题展开的讨论。那时 AI 正在升温，无数有着各种背景、能力和水平的人都在尝试接触 AI 方面的内容，但对于人才的选拔和录用，却似乎没有一个行业内的公认标准和规范。力柯兄的面试题十分简单粗暴，要求面试者仅使用一些基础的 Python 库去实现一个深度神经网络。这听起来有点让人匪夷所思，但事后细想，却是大道至简。这可以让人抛去繁杂的模型，回归神经网络最本质的前向传播和反向传播，将一切都落实在代码层面。虽然需要运用的数学知识不过是一点高等数

学的皮毛，却可以同时从工程和数学两个角度考察候选人的基本功。这几年间，机器学习和深度学习教程及相关公开课越来越多，我阅课无数，竟发现很少有一门课能够沉下心来，仔仔细细地告诉读者和学员，搭建和实现这些神经网络的基础元素从何而来，又为何如此。而本书的上篇，尤其是在第 2 章中，一丝不苟地介绍了神经元、激活函数和损失函数，从偏微分方程层面严谨地推导反向传播，又从代码层面给出了那道面试题的答案。这都让我不由得敬佩力柯兄在工程上的执着。本书的下篇，则是标准的深度学习入门。

至此，我不再“剧透”，因为当你从实战角度阅读这些章节时，会有一种不断发现珍宝的惊喜感，而我更愿意把这些“珍宝”留给本书的读者。

周竟舸，Pinterest 机器学习平台技术负责人

2019 年 12 月

推荐序二

近十余年，计算机领域中令人瞩目的亮点就是以深度学习为代表的一系列突破。无论是人脸检测还是图像识别，抑或文本翻译或无人驾驶，这些在过去几十年里让计算机科学家苦苦思索却不能解的种种难题，在深度学习的帮助下，竟被一一攻克，这不能不说是人类科技史上一颗耀眼的明星。

AI 技术的突飞猛进，却使传统程序员产生不少困惑：过去常用的数据结构、排序搜索、链表数组等，现在变成了模型、卷积、权重和激活函数……无论是要开发 AI 应用，还是和算法研究人员共同工作，他们都存在同一个问题：如何学习 AI 技术？如何理解 AI 算法人员常用的名词和概念？更重要的是，如何把 AI 相关的代码和自己的软件开发经验联系起来？

现在市面上已经有很多深度学习和机器学习教程了，其中也不乏从实例入手、以代码实现为重点的书籍，但并没有一本书真正地从程序员的视角来看待深度学习技术。或者，我们也可以这么说，大部分相关书籍的重点是讲解深度学习理论，所用的实例是解释深度学习理论的实际应用。尽管有不少书籍在讲解理论和代码时详尽而深入，却没有涉及核心问题：要解决这个问题，为什么非用深度学习或机器学习不可？没有这些方法就不能做吗？用深度学习处理该问题的优势是什么？是十全十美，还是存在问题？

打开本书，我惊喜地发现它并非像市面上的其他书籍那样，直接把各种新鲜概念放到读者面前并强迫他们接受。它一开始就没有在机器学习概念上过多纠缠，而是先快速展示了简短的 AI 实现代码的结构和流程，然后带出一些常常让初学者疑惑的问题，针对这些问题再带出新的内容。我们可以看到，本书每一章都用到了类似的形式：阐述一个领域中的实际问题，提出不同的解决方法，简要探讨不同的方法，找到人们难以解决的问题，然后解释机器学习或深度学习处理这些问题的原理。读者了解到的并非单纯的机器学习理论，而是不同领域的具体技术挑战和相关算法的解决方案，从而理解机器学习的真正意义。

必须要说的是，作者在美国工作多年，养成了求真务实和独立思考的习惯，我们从

书中能感受到他独特的风格，并有愉悦的阅读体验。本书在理论讲解方面也没有概念堆砌的枯燥无味，作者常常加入一些对技术的调侃和个人见解，以供读者思考。在代码解析阶段，作者着眼于整体框架与流程，把重点放在理论中的网络结构如何在实际代码中实现，而不会浪费篇幅在代码的语言细节上。

阅读本书，不但是对不同领域 AI 开发的学习，也是一次以资深程序员的视角去审视相关代码实现的体验。本书无论是对于应用开发程序员，还是对于算法研究人员，都相当有价值，非常值得阅读。

喻杰博士，华为智能车云首席技术官

2019 年 12 月

推荐序三

随着 AlphaGo 在人机大战中一举成名，关于机器学习的研究开始备受人们关注。机器学习和神经网络已经被广泛应用于互联网的各个方面，例如搜索、广告、无人驾驶、智能家居，等等。AI 井喷式发展的主要动力如下。

其一，数据的积累。各大 IT 公司都拥有了自己的数据平台，数据积累的速度越来越快。各大高校针对不同的机器学习任务，积累了多样化的数据集。

其二，计算机性能指数级的增长。从当初的 CPU 到 GPU，再从 GPU 到专门为 AI 设计的芯片，都提供了强大而高效的并行计算能力，大大推动了 AI 算法的进步。

其三，AI 理论及模型的突破，例如卷积网络、长短期记忆等。

其四，深度学习开源框架日趋完善。TensorFlow 是当前领先的深度学习开源框架，越来越多的人在使用它从事计算机视觉、自然语言处理、语音识别和一般性的预测分析工作。TensorFlow 集成的 Keras 是为人类而非机器设计的 API，易于学习和使用。

这是一本非常适合程序员入门和实践深度学习的书，理论和实践并重，使用 Keras 作为机器学习框架，侧重于 AI 算法实现。

本书以从代码出发，再回归 AI 相关原理为宗旨，深入浅出、循序渐进地讲解了 Keras 及常见的深度学习网络，还讲解了深度学习在不同领域的应用及模型的部署与服务。读者在一步步探索 AI 算法奥秘的同时，也在享受解决问题的喜悦和成就感，并开启深度学习之旅。

衷心地希望有志于 AI 学习的读者抓住机会，早做准备，成为 AI 时代的弄潮儿。

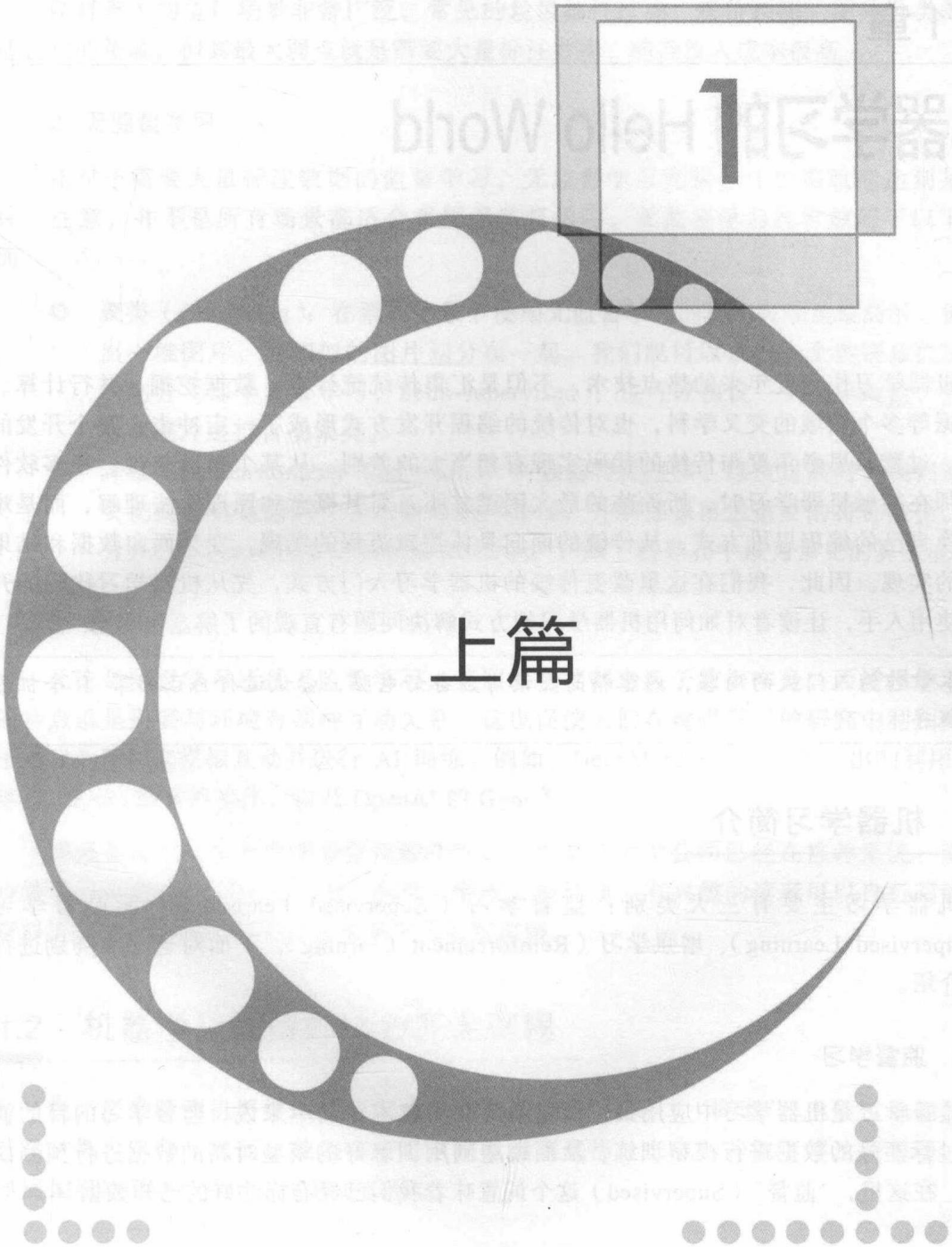
王昀绩，Google AI 高级研究员

2019 年 12 月

目 录

上篇

下篇



1

上篇

第 1 章

机器学习的 Hello World

机器学习作为近年来的热点技术，不但是汇集传统统计学、数据挖掘、并行计算、大数据等多个领域的交叉学科，也对传统的编程开发方式形成了一定冲击，整个开发的模式、过程及思考角度与传统的代码实现有相当大的差别。从某个角度来说，很多软件工程师在接触机器学习时，所面临的最大困难并不是对其概念和原理无法理解，而是难以转换自己的编程思维方式，从传统的面向具体逻辑流程的实现，变为面向数据和结果拟合的实现。因此，我们在这里改变传统的机器学习入门方式，先从机器学习代码的开发和使用入手，让读者对如何用机器学习的方式解决问题有直观的了解。

本章作为入门级的内容，内容精简。读者应备好电脑，尝试运行自己的第 1 个机器学习程序。

1.1 机器学习简介

机器学习主要有三大类别：监督学习（Supervised Learning）、无监督学习（Unsupervised Learning）、增强学习（Reinforcement Learning），下面对这三大类别进行简要介绍。

1. 监督学习

监督学习是机器学习中应用最广泛也最可靠的技术。简单来说，监督学习的目的就是通过标注好的数据进行模型训练，从而期望利用训练好的模型对新的数据进行预测或分类。在这里，“监督”（Supervised）这个词意味着我们已经有标注好的已知数据集。

监督学习的应用场景非常广泛，常见的垃圾邮件过滤、房价预测、图片分类等都是适合它的领域，但其最大弱点就是需要大量标注数据，前期投入成本极高。

2. 无监督学习

相对于需要大量标注数据的监督学习，无监督学习无须标注数据就能达到某个目标。注意，并不是所有场景都适合采用无监督学习，无监督学习经常被用于以下两方面。

- ◎ 聚类（Clustering）：在聚类场景下使用无监督学习的频率可能是最高的。例如给出一堆图片，把相似的图片划分在一起。我们既可以预设一个类别总数进行自动划分（即半监督学习，Semi-supervised）；也可以预设一个差异阈值，然后对所有图片进行自动聚类。
- ◎ 降维（Dimensionality Reduction）：在数据特征过多、维度过高时，我们通常需要把高维数据降到合理的低维空间处理，并期望保留最重要的特征数据。主成分分析（Principal Component Analysis，PCA）就是其中最为常见的算法应用。

3. 增强学习

无论是监督学习还是无监督学习，其训练基础都来源于数据本身。而增强学习最大的特点就是需要与环境有某种互动关系，这也促使人们在增强学习的研究中利用类似电子游戏的环境来模拟互动并进行 AI 训练。例如，DeepMind 在 2015 年提出的利用 DQN 学习 ATARI 游戏的操作，以及 OpenAI 的 Gym 等。

增强学习的实现和应用场景比较特殊，尽管某些大型公司已经在推荐系统、动态定价等场景中尝试应用增强学习，但仍只限于实验性质，有兴趣的读者可以自行阅读其他资料进行学习，在本书中不对增强学习进行讲解。

1.2　机器学习应用的核心开发流程

我们经常听到机器学习的研究人员开口“特征”，闭口“模型”，也听过他们调侃自己是“调参”师，然而，他们口中的这些术语到底指什么呢？若想了解这些术语，就先要清楚机器学习应用开发的核心流程。