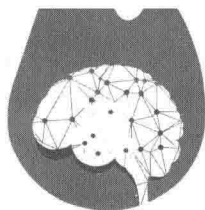


# 深度学习框架 PyTorch 快速开发与实战

邢梦来 王硕 孙洋洋 编著

# 深度学习框架 PyTorch 快速开发与实战

邢梦来 王硕 孙洋洋



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

## 内 容 简 介

深度学习已经成为人工智能炙手可热的技术，PyTorch 是一个较新的、容易上手的深度学习开源框架，目前已得到广泛应用。本书从 PyTorch 框架结构出发，通过案例主要介绍了线性回归、逻辑回归、前馈神经网络、卷积神经网络、循环神经网络、自编码模型、以及生成对抗网络。本书作为深度学习的入门教材，省略了大量的数学模型推导，适合深度学习初学者，人工智能领域的从业者，以及深度学习感兴趣的人阅读。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。  
版权所有，侵权必究。

### 图书在版编目（CIP）数据

深度学习框架 PyTorch 快速开发与实战 / 邢梦来，王硕，孙洋洋编著. —北京：电子工业出版社，2018.8

ISBN 978-7-121-34564-7

I. ①深… II. ①邢… ②王… ③孙… III. ①机器学习 IV. ①TP181

中国版本图书馆 CIP 数据核字（2018）第 135158 号

策划编辑：黄爱萍

责任编辑：张彦红

印 刷：北京季蜂印刷有限公司

装 订：北京季蜂印刷有限公司

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：720×1000 1/16 印张：14.5 字数：232 千字

版 次：2018 年 8 月第 1 版

印 次：2018 年 8 月第 1 次印刷

定 价：69.00 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 [zltz@phei.com.cn](mailto:zltz@phei.com.cn)，盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

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

# 前 言

日常生活中，人工智能悄悄地影响着我们。

随着人工智能的技术的发展，现代几乎各种技术的发展都涉及了人工智能技术，可以说人工智能已经广泛应用到许多领域，其典型的应用包括：信息检索应用、推荐系统、语音识别、自然语言处理、图像识别、智能家居等。以人工智能在语音识别，语音合成上的结果看，2016年10月份由微软美国研究院发布的一个语音识别的最新结果实现了错误率为5.9%的新突破，这是第一次用人工智能技术取得了跟人类似的语音识别的错误率。

人工智能一直处于计算机技术的前沿，人工智能研究的理论和发现在很大程度上将决定计算机技术的发展方向。为了适应新一轮的科技发展，培养高端人才，人工智能进入国家发展战略。

2017年7月，国务院印发《新一代人工智能发展规划》，其中提到，新一代人工智能发展分三步走的战略目标，到2030年使中国人工智能理论、技术与应用总体达到世界领先水平，成为世界主要人工智能创新中心。

为此，我们积极学习人工智能前沿知识，适应科技进步。

本书选用 Facebook 开源深度学习库 PyTorch 作为深度学习框架。常用的深度学习开源平台有 TensorFlow、Theano、Keras、Caffe 等。在 TensorFlow 的官网上，它被定义为一个用于机器智能的开源软件库，使用 TensorFlow 需要编写大量的代码，个人觉得不适合初学者。

Theano 是比较老牌和最稳定的库之一。由于 Theano 不支持多 GPU 扩展，在深度学习开源平台快速更新迭代的浪潮下，Theano 已然开始慢慢被遗忘了。

Keras 句法比较明晰，文档完善，使用非常简单轻松。Keras 强调极简主义，只需几行代码就能构建一个神经网络，适合新人学习。

Caffe 是老牌中的老牌框架。起初的时候它仅仅关注计算机视觉，但它具有非常好的通用性。Caffe 的缺点是它不够灵活，同时 Caffe 的文档非常贫乏。

张量是 PyTorch 的一个完美组件，和 NumPy 类似。将张量从 NumPy 转换至 PyTorch 非常容易。可以把它作为 NumPy 的替代品。PyTorch 这种框架可以获得 GPU 加速，以便快速进行数据预处理，或其他任务。PyTorch 同时也提供了变量，我们在构建神经网络的时候，在张量之上的封装，构建自己的计算图，并自动计算梯度。PyTorch 建立的是动态图，TensorFlow 建立的是静态图。PyTorch 更加符合一般的编程习惯，而不是像 TensorFlow 那样需要先定义计算图。

虽然开源平台众多，但更多的时候，我们考虑实现算法的简捷性，通常选择容易上手的，能快速实现算法的开源平台。为此，我们需要选择适合自己的深度学习开源平台，实现深度学习算法。

学习深度学习理论知识，了解人工智能行业发展动态，掌握前沿科学技术。利用 PyTorch 开源平台快速实现经典卷积神经网络、循环神经网络、自编码模型、对抗生成网络等模型。开启海绵模式，尽可能多学原理知识，掌握机器学习的基础理论知识，然后针对性地训练。通常从收集数据，预处理和清洗数据，到搭建模型，训练和调试模型，再到最后评估模型。逐渐培养出对于什么样的数据适合用什么类型的模型的判断能力，并增强实践能力。经过学习，逐渐从“小白”，慢慢到专业人士。有兴趣的读者欢迎加入本书交流群，一起交流学习。同时，本书所有案例的代码统一放在 QQ 群文件里，群号为 662443475，或在博文视点官网下载：[www.broadview.com.cn](http://www.broadview.com.cn)。

## 致谢

感谢电子工业出版社的黄爱萍编辑，在选题策划和稿件整理方面做出的大量工作。

感谢极宽开源量化团队给予的技术支持。

在本书的创作中，特别感谢张建辉、刘笑俐、王丽颖、刘晓峰、刘婷、沈雨涵的协助，为他们的付出表示感谢。

邢梦来

2018年6月

轻松注册成为博文视点社区用户 ([www.broadview.com.cn](http://www.broadview.com.cn))，扫码直达本书页面。

- **下载资源：**本书如提供示例代码及资源文件，均可在 [下载资源](#) 处下载。
- **提交勘误：**您对书中内容的修改意见可在 [提交勘误](#) 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动：**在页面下方 [读者评论](#) 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/34564>



# 目 录

## 第一部分 理论部分

第 1 章 深度学习简介	2
1.1 深度学习	2
1.2 神经网络的发展	6
1.3 深度学习的应用	7
1.4 常用的数学知识和机器学习算法	8
1.5 PyTorch 简介	11
1.5.1 PyTorch 介绍	11
1.5.2 使用 PyTorch 的公司	15
1.5.3 PyTorch API	16
1.5.4 为什么选择 Python 语言	16
1.5.5 Python 语言的特点	16
1.6 常用的机器学习、深度学习开源框架	17
1.7 其他常用的模块库	19
1.8 深度学习常用名词	20
第 2 章 PyTorch 环境安装	33
2.1 基于 Ubuntu 环境的安装	33
2.1.1 安装 Anaconda	35

2.1.2 设置国内镜像	36
2.2 Conda 命令安装 PyTorch	37
2.3 pip 命令安装 PyTorch	37
2.4 配置 CUDA	38
第 3 章 PyTorch 基础知识	40
3.1 张量	40
3.2 数学操作	43
3.3 数理统计	44
3.4 比较操作	45
第 4 章 简单案例入门	47
4.1 线性回归	47
4.2 逻辑回归	52
第 5 章 前馈神经网络	59
5.1 实现前馈神经网络	61
5.2 数据集	68
5.3 卷积层	72
5.4 Functional 函数	75
5.5 优化算法	82
5.6 自动求导机制	85
5.7 保存和加载模型	87
5.8 GPU 加速运算	87
第 6 章 PyTorch 可视化工具	89
6.1 Visdom 介绍	89
6.2 Visdom 基本概念	90
6.2.1 Panes (窗格)	90
6.2.2 Environments (环境)	90
6.2.3 State (状态)	91
6.3 安装 Visdom	91
6.4 可视化接口	91



6.4.1	Python 函数属性提取技巧	92
6.4.2	vis.text	93
6.4.3	vis.image	93
6.4.4	vis.scatter	94
6.4.5	vis.line	95
6.4.6	vis.stem	97
6.4.7	vis.heatmap	97
6.4.8	vis.bar	99
6.4.9	vis.histogram	101
6.4.10	vis.boxplot	102
6.4.11	vis.surf	103
6.4.12	vis.contour	104
6.4.13	vis.mesh	106
6.4.14	vis.svg	107

## 第二部分 实战部分

第 7 章	卷积神经网络	110
7.1	卷积层	112
7.2	池化层	114
7.3	经典的卷积神经网络	115
7.3.1	LeNet-5 神经网络结构	115
7.3.2	ImageNet-2010 网络结构	117
7.3.3	VGGNet 网络结构	122
7.3.4	GoodLeNet 网络结构	124
7.3.5	ResNet 网络结构	126
7.4	卷积神经网络案例	129
7.5	深度残差模型案例	138
第 8 章	循环神经网络简介	145
8.1	循环神经网络模型结构	146
8.2	不同类型的 RNN	147
8.3	LSTM 结构具体解析	151

8.4	LSTM 的变体	153
8.5	循环神经网络实现	156
8.5.1	循环神经网络案例	156
8.5.2	双向 RNN 案例	160
第 9 章	自编码模型	164
第 10 章	对抗生成网络	172
10.1	DCGAN 原理	175
10.2	GAN 对抗生成网络实例	180
第 11 章	Seq2seq 自然语言处理	186
11.1	Seq2seq 自然语言处理简介	186
11.2	Seq2seq 自然语言处理案例	188
第 12 章	利用 PyTorch 实现量化交易	204
12.1	线性回归预测股价	205
12.2	前馈神经网络预测股价	209
12.3	递归神经网络预测股价	214

# 1

## 第一部分

---

### 理论部分

- 第 1 章 深度学习简介
- 第 2 章 PyTorch 环境安装
- 第 3 章 PyTorch 基础知识
- 第 4 章 简单案例入门
- 第 5 章 前馈神经网络
- 第 6 章 PyTorch 可视化工具

# 1

## 第 1 章

# 深度学习简介

在北京时间 2016 年 3 月 15 日的下午，谷歌 DeepMind 团队开发的围棋深度学习系统 AlphaGo 以总比分 4 : 1 战胜了韩国棋手李世石，成为第一个在 19×19 棋盘上战胜人类围棋冠军的智能系统。2017 年 10 月 19 日，DeepMind 团队重磅发布 AlphaGo Zero。相比上一代 AlphaGo，该版本的 AlphaGo 实现了在 AI 发展中非常有意义的一步——“无师自通”。在 AlphaGo 的核心组成部分中，估值网络 (Value Network) 和走棋网络 (Policy Network) 都使用了深度学习的技术。AlphaGo 战胜李世石的新闻成功地把深度学习的概念从学术界推向了大众，并点燃了大众对于人工智能的巨大热情。AlphaGo Zero 的伟大之处是第一次让机器可以不通过任何棋谱，不通过任何人类的经验，在只告诉规则的前提下，成为一个围棋高手，这种无师自通的学习模式在 AI 整个发展历史上是非常有意义的。

### 1.1 深度学习

深度学习的概念由 Hinton 等人于 2006 年提出。深度学习的概念源于神经网络的研究。含多隐层的多层感知器就是一种深度学习结构。深度学习通过组合低层特征形成更加抽象的高层次表示属性类别或特征，以发现数据的分布式特征表示。

深度学习在完成一些难度极高的任务中展现了惊人的功力，如从图片中识别物体、语言理解，棋盘类游戏等。我们举一个例子来具体说明什么是深度学习。

我们都知道机器学习的目的是在没有特定编程的情况下，希望系统去回答某个问题。比如：明天北京会下雨吗？这类问题可以翻译成以下形式：对于给定的输入  $X$ ，正确的输出  $Y$  是什么？输入的是北京天气的信息，输出则为会下雨或者不会下雨。

传统的机器学习需要我们首先定义一系列程序寻找的特征。比如，图 1.1 里的猫有两个眼睛一个鼻子，四条腿，以及它毛茸茸的毛。我们给程序提供大量的例子，但是对于每个例子，都不会向程序展示全局图片，而是展示预先设定的某特征的变量，然后告诉程序哪个才是猫。经过训练之后，程序就能明白如果它没有毛茸茸的毛，它可能不是猫。而深度学习却可以很好地解决这个问题。

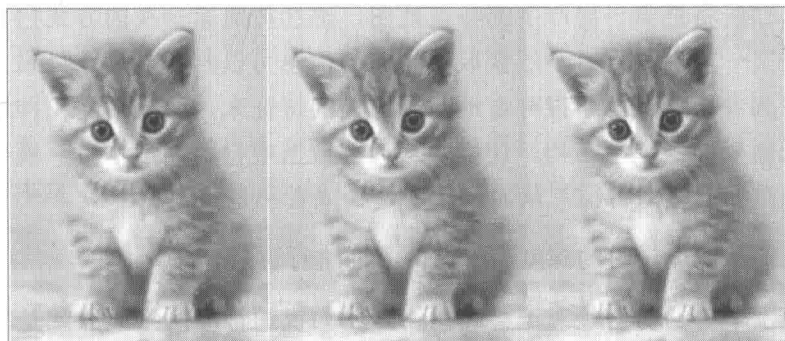


图 1.1 使用深度学习识别图片中的猫

深度学习模型中的卷积神经网络模型可以用来解决这个问题，通过卷积神经网络我们来讲讲深度学习模型是如何识别图片里的猫的。图 1.2 是简单的单层神经网络模型。

第一步，我们准备好要识别的图片，假设这张图片的尺寸为  $28 \times 28$  像素，然后把图片中的像素传输给卷积神经网络模型。第一层神经网络将扫描图片，以  $5 \times 5$  的色块为单位，去寻找一些基本的特征，并提取特征，形成特征地图。

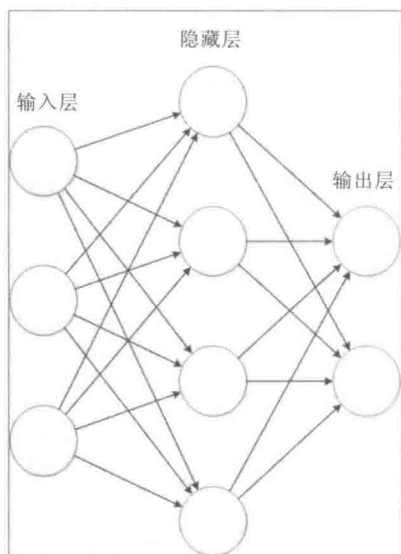


图 1.2 单层神经网络模型

第二层将把第一层产生的特征地图，继续按照第一次的方式扫描，这样一层又一层，直到有一层获取了足够的信息可以判断这是一只猫。我们可以看到卷积神经网络很神奇地筛选出猫的特征来，而采用卷积神经网络我们只需要设定模型结构，让程序从训练数据中自己归纳学习，模型会自动采集重要的模型特征，从而判断出是猫或者不是猫。

我们从深度学习的深度和学习两方面进行讲解。

深度学习的前身是人工神经网络。我们说的神经网络一般就是指人工神经网络。最简单的神经网络由输入层、隐藏层和输出层组成。输入训练数据为输入层，输出计算结果为输出层，隐藏层使输入数据传播到输出层，从而把神经网络形成网络结构。传统神经网络的每一层有大量的节点组成，也叫神经元。每层内的节点相互独立、互不干扰。层与层之间的节点相互连接，深度学习就是增加多层网络结构，利用现有的数据，来对未知的数据做预测分类。

深度学习的模型是如何学习的呢？我们也许都有过这样的经历，做过一道数学题并且知道如何求解，下次再遇到类似的题目就可以很轻松地解决。我们在学习的过程中举一反三、触类旁通，逐渐地拥有解决问题的能

力。但是我们都知机器没有大脑思维，如何才能让机器自己学习并拥有这样的能力呢？这也是人工智能发展的方向。现在我们会问深度学习是否有这样的能力呢？

我们给设定好的机器模型输入大量的数据，不断地训练，最终得到我们想要的结果。这些过程就像我们平时做大量练习题，不断地思考并且训练如何能正确地解答题目，到最后获得正确的解决问题、得出答案的能力。

我们再从监督学习和无监督学习两方面进行讲解。

机器学习方法有监督学习与无监督学习之分，和机器学习方法一样，深度机器学习方法也有监督学习与无监督学习之分。不同的学习框架下建立的学习模型也不同。

有监督学习：利用一组已知类别的样本调整分类器的参数，使其达到所要求性能的过程，也称为监督训练或有教师学习。对具有概念标记（分类）的训练样本进行学习，以尽可能对训练样本集外的数据进行标记（分类）预测。这里，所有的标记（分类）是已知的。因此，训练样本的歧义性低。

无监督学习：对没有概念标记（分类）的训练样本进行学习，以发现训练样本集中的结构性知识。这里，所有的标记（分类）是未知的。因此，训练样本的歧义性高。聚类就是典型的无监督学习。

半监督学习（Semi-supervised Learning）是模式识别和机器学习领域研究的重点问题，是监督学习与无监督学习相结合的一种学习方法。它主要考虑如何利用少量的标注样本和大量的未标注样本进行训练和分类的问题。半监督学习对于减少标注代价，提高机器学习性能具有非常重大的意义。

虽然 AlphaGo 战胜李世石将人工智能推向了一个新的高度，但是 AlphaGo 能够解决的仅仅是在一个特定环境中定义好的问题，要将人工智能系统真正的应用到开放环境，还需要研究人员更多的努力。这也将是 AI 未来发展的方向。

## 1.2 神经网络的发展

2017年，人工智能发展火热。机器学习和大数据领域取得突破性的进展，作为人工智能一个重要分支的深度学习，也正受到大家越来越多的关注。经历了几年的高速发展，深度学习在工业界和学术界备受追捧，百度、阿里巴巴、腾讯也将人工智能方面的人才培养作为战略重心，同时吹响了人才抢夺的号角。斯坦福大学教授、计算机视觉领域领军人物李飞飞(Feifei Li)于2016年加入谷歌；卡内基梅隆大学教授、机器学习领域顶级人物Alex Smola于2016年加入亚马逊。深度学习界泰斗吴恩达(Andrew Ng)宣布退出百度，成立了自己的人工智能公司，2017年8月15日，吴恩达向美国证券交易委员会(SEC)注册了一支1.5亿美元的风险投资基金，专门投资人工智能领域。

除吴恩达外，在过去的几个月中，谷歌也公布了一项专注人工智能领域的风险投资基金——Gradient Ventures，它将为人工智能领域的初创公司提供科研资金和技术指导。Dropbox前创始人和腾讯也一起创立了1.36亿美元的Basis Set Ventures，同样专注人工智能领域。此外，Element.AI筹集了1.02亿美元；微软创立了自己的人工智能基金；丰田也筹备了1亿美元用于人工智能投资。

深度学习为什么会这么火？我们也许可以从生活中找到答案。比如阿里巴巴在杭州的无人超市，百度正在研发的无人驾驶的汽车，人脸识别软件等，这些都是深度学习在生活中的应用。也许很多人都觉得奇怪，看似没有任何情感的机器如何能完成一些复杂的操作呢？

下面开始介绍深度学习的发展历程，也许会找到相关答案。

深度学习属于机器学习一个分支，同时机器学习又是人工智能的一个子集。什么叫作人工智能呢？在20世纪40年代和50年代，来自不同领域的一批科学家开始探讨制造人工大脑的可能性。1956年，人工智能被确立为一门学科。人工智能是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。相信大家都听说过“图灵测试”吧？图灵测试一词来源于计算机科学和密码学的先驱阿兰·麦席森·图灵写于1950年的一篇论文《计算机器与智能》。阿兰·麦席森·图



灵 1950 年设计出这个测试，其内容是，如果电脑能在 5 分钟内回答由人类测试者提出的一系列问题，且其超过 30% 的回答让测试者误认为是人类所答，则电脑通过测试。如果电脑通过测试，则说明智能程度相对较高。为此，科学家正在研究如何让机器像人类一样思考。

什么叫作机器学习呢？机器学习（Machine Learning, ML）是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。它专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。它是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域，它主要使用归纳、综合而不是演绎的方法。

## 1.3 深度学习的应用

2017 年 3 月 13 日，AlphaGo 与李世石的第四场对决结束，在连输 3 场之后，李世石终于扳回一局。但 3:1 的比赛结果已说明人工智能的强大，这也是谷歌对深度学习、人工智能的成功营销。AlphaGo 在短短几个月实现性能的大幅提升，用五个月走完了 IBM “深蓝” 4 年的路，体现了当前人工智能系统学习速度之快。

2017 年 10 月 30 日，MIT Technology Review、Slate、Quartz、Gear of Biz 等美国媒体发表文章称，用不了多长时间，AlphaGo 将不再是地球上最好的棋手。之后新式高超的人工智能程序版本 AlphaGo Zero 出现，它堪称怪物。它从零开始，面对的只是一张空白的棋盘和游戏规则。它无师自通，仅仅通过自学使自己的游戏技能得以提高。但是它从来都不仅仅关乎棋盘游戏，未来它将会在更多领域发挥作用。

近年来，深度学习已经在图像识别、语音识别等领域获得了一些应用。目前深度学习技术应用最多的还是视觉领域，即对图像和视频的分析。在图像分析方面，比如人们熟悉的人脸识别、文字识别和大规模图像分类等。深度学习大幅提升了复杂任务分类的准确率，使得图像识别、语音识别以及语义理解准确率大幅提升。在语音识别等领域应用有 iPhone 的语音助理 Siri、百度的度秘、科大讯飞的“灵犀”、微软的小冰等。