

深度学习 原理与实践

THE PRINCIPLE AND PRACTICE OF DEEP LEARNING

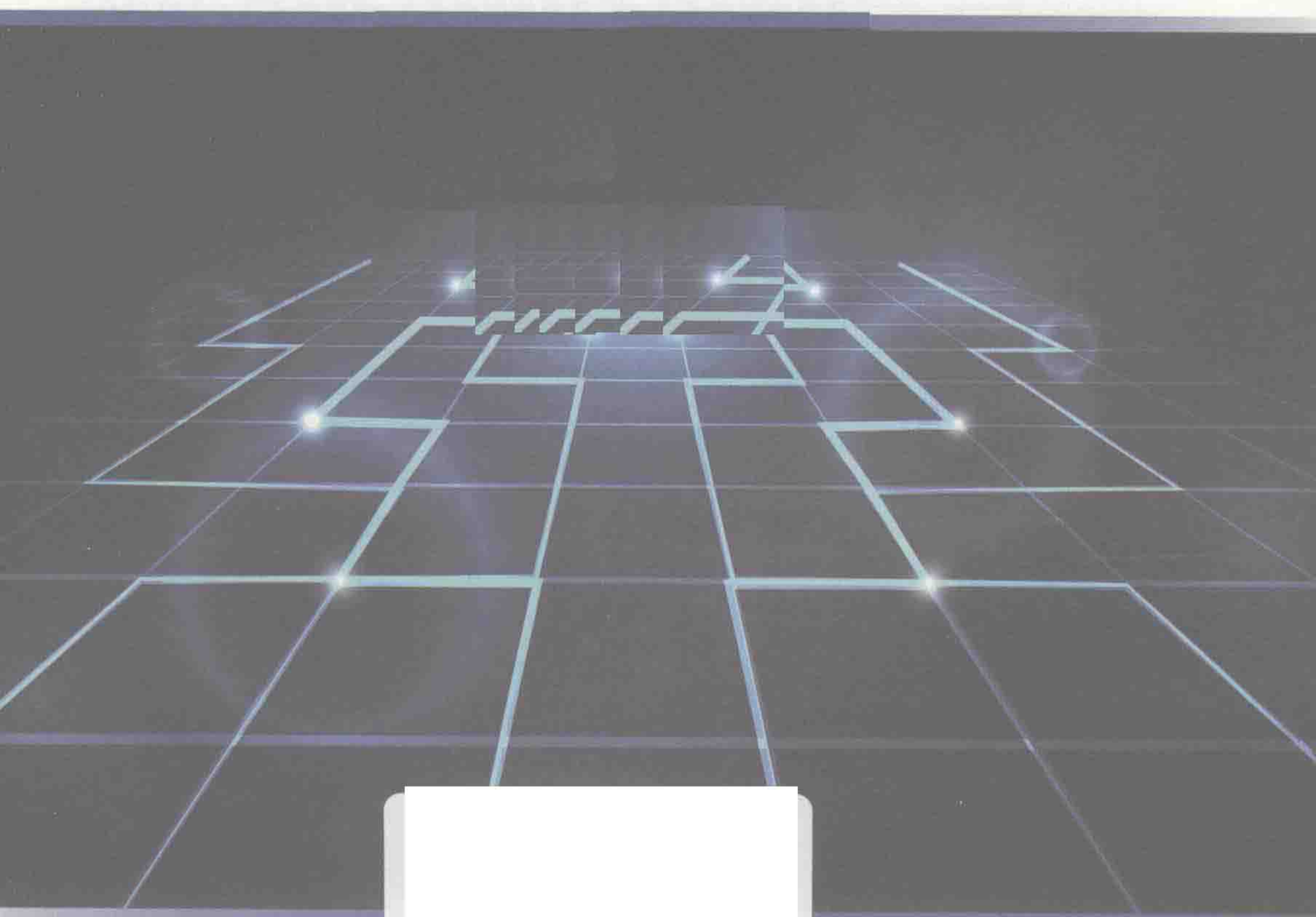
陈仲铭 彭凌西◎著

非外借

深度学习 原理与实践

THE PRINCIPLE AND PRACTICE OF DEEP LEARNING

陈仲铭 彭凌西◎著



人民邮电出版社

北京

图书在版编目 (C I P) 数据

深度学习原理与实践 / 陈仲铭, 彭凌西著. -- 北京:
人民邮电出版社, 2018. 8
ISBN 978-7-115-48367-6

I. ①深… II. ①陈… ②彭… III. ①机器学习—研
究 IV. ①TP181

中国版本图书馆CIP数据核字(2018)第112818号

内 容 提 要

本书详细介绍了目前深度学习相关的常用网络模型 (ANN、CNN、RNN), 以及不同网络模型的算法原理和核心思想。本书利用大量的实例代码对网络模型进行了分析, 这些案例能够加深读者对网络模型的认识。此外, 本书还提供完整的进阶内容和对应案例, 让读者全面深入地了解深度学习的知识和技巧, 达到学以致用目的。

本书适用于大数据平台系统工程师、算法工程师、数据科学家, 可作为对人工智能和深度学习感兴趣的计算机相关从业人员的学习用书, 也可作为计算机等相关专业的师生用书和培训学校的教材。

-
- ◆ 著 陈仲铭 彭凌西
责任编辑 张 爽
责任印制 焦志炜
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京瑞禾彩色印刷有限公司印刷
 - ◆ 开本: 720×960 1/16
印张: 21.25
字数: 392 千字
印数: 1-3 000 册
- 2018 年 8 月第 1 版
2018 年 8 月北京第 1 次印刷
-

定价: 89.00 元

读者服务热线: (010) 81055410 印装质量热线: (010) 81055316

反盗版热线: (010) 81055315

广告经营许可证: 京东工商广登字 20170147 号

序

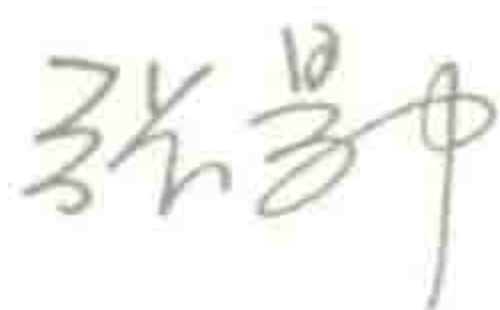
人工智能领域一直备受国内外著名研究专家和学者的关注。近年来，以机器学习、深度学习、神经网络及模拟计算为代表的计算智能技术得到空前发展，掀起了人工智能领域的新高潮。

这其中特别值得一提的是深度学习。科学家已经证明，深度学习技术的深度表征能力能够有效地提取数据的高维特征。深度学习模型从简单的特征开始，通过神经网络逐层组合的方式，不断地抽取更加复杂的特征到下一层，最终提取到高维的抽象特征。深度学习模型具有特殊的网络结构组织方式，使网络层数不断加深，加上特殊模型训练技巧，反向梯度算法和随机梯度算法的组合，在理论上可以表达任何函数的网络模型。深度学习因为具有的大规模并行分布式处理、自组织、自学习、自适应、可迁移学习的能力，以及数据多样性，对序列信号有记忆功能和鲁棒性等特点，因此受到国内外众多学者的高度重视。

受深度学习这一技术的影响，近年来涌现出了 ImageNet、COCO、Pascal、VOT 等与深度学习相关的全球算法竞赛，腾讯、阿里巴巴、Google、Facebook、Amazon 等国内外著名互联网企业相继建立各自的人工智能实验室。仅 2017 年，国际上举办有关深度学习的著名专题会议高达 28 场，其中包括 ICCV、ICRL、NIPS 等。

基于深度学习与生物神经网络的相似性，近年来研究者们提出了深度神经网络这一新概念。在计算机推荐系统、图像感知、模拟学习等领域，传统机器学习的弱点是其表征能力受限。深度学习的及时出现，为众多应用领域带来新的挑战，具有广阔的应用前景。

“深度学习”的相关图书近年来层出不穷，但是将原理阐述和实践紧密结合的书籍相对较少。本书对深度学习的基本概念、原理及应用技术做了深入浅出的讲解，它的出版对从事人工智能、机器学习、深度学习研究的科技工作者和研究生大有裨益。相信本书的出版会促进国内深度学习的研究和应用，有助于我国新一代人工智能创新活动的蓬勃发展。



中国科学院院士

2018 年 1 月 10 日于广州

前言

随着深度学习的爆发，截至目前，最新的深度学习模型已经远远超越了传统的机器学习算法，在数据的预测准确性和分类精度上都达到了前所未有的高度。深度学习的表征功能不需要工程师手动提取数据中的隐含特征，而是自动地深度挖掘数据并提取数据中隐含的高维特征。可以说深度学习实实在在地节省了工程师们在特征工程上的时间，提高了算法的精度和准确率，也推动着实际工程的进步。

深度学习的出现为机器学习和人工智能领域带来了新的机会和希望，极大地拓展了人们对于计算人工智能的想象力和应用范畴，使很多的计算辅助功能成为可能。深度学习已经不仅仅是计算机科学领域的问题，它结合了神经科学和逻辑学科的知识，涉及众多跨学科领域的知识交叉。从图像感知、无人驾驶汽车、无人驾驶飞机，到生物医学的预防性诊断、病理预测，甚至是更加贴近年轻一代的电影推荐、购物指南，几乎所有领域都可以使用深度学习。而在某些领域中，人类所掌握的知识还是如此地微不足道，因此深度学习在这些领域仍有巨大的发展空间。

在深度学习领域最权威的书是 Ian Goodfellow、Yoshua Bengio 和 Aaron Courville 编写的 *Deep Learning*（中文版由人民邮电出版社出版）。但与深度学习相关，且将知识原理与案例实践相结合的书籍在市场上并不多见。作者写作本书的初衷就是帮助更多的人了解深度学习，并投身于人工智能领域。

一年多以来，作者业余时间几乎都用于写作、编写示例代码，也曾经想过放弃，但都坚持了下来。成功不就是一一次次坚持而成就的吗？人生，多多少少也有些梦想，值得我们去付出！

学习建议

Conquer yourself rather than the world.

征服你自己，而不是去征服世界。

——笛卡儿《谈谈方法》

我们努力提高自己的能力、学习深度学习和人工智能，并不是为了与别人一决高下，而是超越自己。成功的路上会有很多建议，而针对本书的学习，有下面几条建议。

1. 学习 Python 基础知识

贯穿全书中的代码由 Python 编程语言编写而成，但是本书并没有对 Python 代码进行过多的解释，而是围绕和专注于讲解深度学习的原理和思想。因此希望读者先自行安装 Python 3，了解 Python 编程语言的特性和基本使用方法，有一定的了解后即可直接使用本书的示例代码，这对于理解深度学习会有所帮助。

2. 实践是检验真理的唯一标准

每一行代码和函数都是作者亲自实现过的，尽管不同服务器和不同版本的框架会存在一些差异，也可能读者看到本书代码时已经稍显过时，但我仍然希望读者亲自去尝试实现书中的代码，毕竟“纸上得来终觉浅”，实践才是检验真理的唯一标准。

3. 不要纠结于框架

深度学习好比“菜谱”，数据就是“肉和青菜”，深度学习框架就是“炒菜的锅和铲”。谁说红烧土豆一定要用“Caffe”牌的锅，香煎莲藕一定要用“Google”牌的铲？我们只要学会菜谱，研究“锅”和“铲”的属性并多加实践练习，就可以炒出一盘香喷喷的菜。希望读者不要纠结于本书所用的是 Keras 框架，还是 Tensorflow 框架，而是将所有的目光都聚焦在深度学习的原理和案例实践上。

4. 多阅读相关文献

读杰出的书籍，有如和杰出的人物促膝交谈。深度学习的知识日新月异，在知识更新迭代迅速的时代，我们需要掌握知识的本质内容，而掌握本质内容的最好方法之一就是阅读与知识点相关的论文文献，去理解与思考为什么要这样，这样做的优缺点是什么。多阅读相关的文献，我们就会更好地把握住深度学习知识的本质。

本书特色

(1) 大量图例，简单易懂。作者亲自绘制了大量插图，力求还原深度学习的算法思想，分解和剖析晦涩的算法，用图例来表示复杂的问题。生动的图例也能给读者带来阅读乐趣，快乐地学习算法知识，体会深度学习的算法本质。

(2) 简化公式，生动比喻。深度学习和机器学习类的书中通常会有大量复杂冗长的算法公式，为了避免出现读者读不懂的情况，本书尽可能地统一了公式和符号，简化相关公式，并加以生动的比喻进行解析。在启发读者的同时，锻炼读者分析问题和解决问题的能力。

(3) 算法原理，代码实现。在介绍深度学习及相关算法的原理时，不仅给出了对应的公式，还给出了实现和求解公式的代码，让读者明确该算法的作用、输入和输出。原理与代码相结合，使得读者对深度学习的算法实现更加具有亲切感。

(4) 深入浅出，精心剖析。理解深度学习需要一定的机器学习知识，本书在第 1 章介绍了深度学习与机器学习的关系，并简要介绍了机器学习的内容。在内容安排

上，每章依次介绍模型框架的应用场景、结构和使用方式，最后通过真实的案例去全面分析该模型结构。目的是让读者可以抓住深度学习的本质。

(5) 入门实践，案例重现。每一章最后的真实案例不是直接堆砌代码，而是讲解使用该算法模型的原因和好处。从简单的背景知识出发，使用前文讲解过的深度学习知识实现一个实际的工程项目。实践可以用于及时检验读者对所学知识的掌握程度，为读者奠定深度学习的实践基础。

建议和反馈

为了让广大读者更好地理解和使用书中的案例代码，本书提供了一个公开的 GitHub 代码库：<https://github.com/chenzomi12/deeplearning-inaction/>。

完成一本书的编著与出版是一项极其琐碎和繁重的工作，我已尽力完善本书内容，但仍然可能存在纰漏和错误之处，敬请各位读者朋友指正，请致信作者邮箱 chenzomi12@gmail.com 或本书编辑邮箱 zhangshuang@ptpress.com.cn。

作者衷心地希望各位读者能够从本书获益，这是对我最大的支持和鼓励。

致谢

感谢每一位为此书做出贡献的朋友。

感谢每一位读者，希望本书的内容值得您投入宝贵的时间。

感谢王佳博士在本书目录和内容结构上提出的建议。

感谢人民邮电出版社的张爽编辑，感谢您的精心策划和建议。

感谢广东海格通讯有限公司的领导吕韶清和星奥科技的领导古明晖，让我在工作中有机会接触机器视觉、机器学习和深度学习，这是我开始撰写本书的契机。

感谢广东工业大学的研究生刘尚武、中山大学的段永强两位兄弟对本书的大力支持，他们对本书进行了多次的审阅和批注，并提出了宝贵的意见，本书的每一章内容都经过他们两人的精心修改。正是他们的付出，才有了这本通俗易懂的深度学习读物。

感谢我尊敬的父母、亲爱的姐姐陈泳茵以及爱人欧阳素行，在高度紧张的工作氛围和高强度加班的环境下，在我业余时间给予我大力的支持和鼓励，让我有勇气和耐心完成一次又一次的改动和编辑，并在写作的语言细节上给了我很多启发。

最后，感谢国家自然科学基金资助（编号 61100150 和 61772147）、广东省高校创新团队项目资助（编号 2015KCXTD014），以及广东省高等学校自然科学研究重大项目资助（编号 2014KZDXM044）。

陈仲铭

2018年1月

目 录

第1章 初探深度学习	1
1.1 什么是深度学习	2
1.1.1 机器学习的一般方法	4
1.1.2 选择深度学习的原因	5
1.1.3 深度学习前的思考	6
1.2 深度学习的应用	7
1.3 深度学习的硬件加速器	10
1.3.1 GPU比CPU更适合深度学习	10
1.3.2 GPU硬件选择	13
1.4 深度学习的软件框架	15
1.5 本章小结	19
引用/参考	19
第2章 人工神经网络	21
2.1 人工神经网络概述	22
2.1.1 历史背景	22
2.1.2 基本单位——神经元	24
2.1.3 线性模型与激活函数	25
2.1.4 多层神经网络	26
2.2 训练与预测	28
2.2.1 训练	28
2.2.2 预测	29
2.3 核心算法	29
2.3.1 梯度下降算法	29
2.3.2 向前传播算法	31

2.3.3	反向传播算法	33
2.4	示例：医疗数据诊断	42
2.4.1	从医疗数据到数学模型	43
2.4.2	准备数据	44
2.4.3	线性分类	45
2.4.4	建立人工神经网络模型	46
2.4.5	隐层节点数对人工神经网络模型的影响	53
2.5	本章小结	55
	引用/参考	56
第3章	深度学习基础及技巧	58
3.1	激活函数	59
3.1.1	线性函数	60
3.1.2	Sigmoid函数	61
3.1.3	双曲正切函数	62
3.1.4	ReLU函数	63
3.1.5	Softmax函数	64
3.1.6	激活函数的选择	65
3.2	损失函数	65
3.2.1	损失函数的定义	66
3.2.2	回归损失函数	67
3.2.3	分类损失函数	69
3.2.4	神经网络中常用的损失函数	70
3.3	超参数	71
3.3.1	学习率	71
3.3.2	动量	72
3.4	深度学习的技巧	73
3.4.1	数据集准备	73
3.4.2	数据集扩展	74
3.4.3	数据预处理	76
3.4.4	网络的初始化	81
3.4.5	网络过度拟合	84
3.4.6	正则化方法	85

3.4.7	GPU的使用	88
3.4.8	训练过程的技巧	89
3.5	本章小结	91
	引用/参考	92
第4章	卷积神经网络	93
4.1	卷积神经网络概述	94
4.1.1	卷积神经网络的应用	95
4.1.2	卷积神经网络的结构	99
4.1.3	卷积神经网络与人工神经网络的联系	101
4.2	卷积操作	103
4.2.1	滑动窗口卷积操作	104
4.2.2	网络卷积层操作	105
4.2.3	矩阵快速卷积	107
4.3	卷积神经网络三大核心思想	110
4.3.1	传统神经网络的缺点	110
4.3.2	局部感知	111
4.3.3	权值共享	112
4.3.4	下采样	113
4.4	设计卷积神经网络架构	115
4.4.1	网络层间排列规律	116
4.4.2	网络参数设计规律	116
4.5	示例1: 可视化手写字体网络特征	117
4.5.1	MNIST手写字体数据库	118
4.5.2	LeNet5网络模型	119
4.5.3	LeNet5网络训练	122
4.5.4	可视化特征向量	124
4.6	示例2: 少样本卷积神经网络分类	127
4.6.1	Kaggle猫狗数据库	128
4.6.2	AlexNet模型	130
4.6.3	AlexNet训练	134
4.6.4	AlexNet预测	136
4.6.5	微调网络	137

4.7	本章小结	140
	引用/参考	141
第5章	卷积神经网络视觉盛宴	143
5.1	图像目标检测	144
5.1.1	传统目标检测方法	146
5.1.2	基于区域的网络	147
5.1.3	基于回归的网络	157
5.1.4	目标检测小结	163
5.2	图像语义分割	165
5.2.1	传统图像分割方法	165
5.2.2	全卷积神经网络	167
5.2.3	SegNet网络	171
5.2.4	DeepLab网络	173
5.3	示例1: NMS确定候选框	176
5.4	示例2: SS算法提取候选框	179
5.4.1	图像复杂度	179
5.4.2	算法核心思想	180
5.4.3	区域相似度计算	184
5.5	本章小结	189
	引用/参考	190
第6章	卷积神经网络进阶示例	192
6.1	示例1: 全卷积网络图像语义分割	193
6.1.1	VGG连续小核卷积层	193
6.1.2	VGG网络模型	195
6.1.3	全卷积网络模型	199
6.1.4	全卷积网络语义分割	204
6.2	示例2: 深度可视化网络	209
6.2.1	梯度上升法	210
6.2.2	可视化所有卷积层	213
6.2.3	可视化输出层	218
6.2.4	卷积神经网络真能理解视觉吗	219

6.3	示例3: 卷积神经网络艺术绘画	220
6.3.1	算法思想	221
6.3.2	图像风格定义	222
6.3.3	图像内容定义	224
6.3.4	算法实现	225
	引用/参考	229
第7章	循环神经网络	231
7.1	初识循环神经网络	232
7.1.1	前馈式神经网络的缺点	233
7.1.2	什么是序列数据	234
7.2	循环神经网络的应用	235
7.3	循环神经网络的模型结构	237
7.3.1	序列数据建模	237
7.3.2	基本结构	238
7.3.3	其他结构	239
7.4	循环神经网络的核心算法	241
7.4.1	模型详解	241
7.4.2	向前传播	243
7.4.3	损失函数	245
7.4.4	时间反向传播算法	246
7.4.5	梯度消失与梯度爆炸	251
7.5	示例: 使用循环神经网络预测文本数据	254
7.5.1	定义网络模型	254
7.5.2	序列数据预处理	255
7.5.3	准备输入输出数据	258
7.5.4	实现简单的循环神经网络模型	260
7.6	本章小结	263
	引用/参考	264
第8章	循环神经网络进阶序列长期记忆	265
8.1	长期依赖问题	266
8.2	长短期记忆网络	269

8.2.1	LSTM网络结构	269
8.2.2	LSTM记忆单元	270
8.2.3	LSTM记忆方式	274
8.3	门控循环单元	274
8.3.1	GRU记忆单元	275
8.3.2	GRU实现	276
8.3.3	GRU与LSTM比较	277
8.4	示例1: 神奇的机器翻译	278
8.4.1	基于统计的机器翻译	279
8.4.2	基于神经网络的机器翻译	282
8.4.3	编码-解码模型	283
8.4.4	平衡语料数据集	287
8.4.5	机器翻译的未来	292
8.5	示例2: 智能对话机器人	293
8.5.1	Seq2Seq模型	294
8.5.2	Seq2Seq模型的缺点	295
8.5.3	超越Seq2Seq框架	296
8.6	示例3: 智能语音识别音箱	299
8.6.1	语音识别框架	300
8.6.2	准备语音数据	302
8.6.3	语音特征提取	306
8.6.4	声学模型	311
8.6.5	语言模型	323
8.6.6	语音识别的展望	323
8.7	本章小结	324
	引用/参考	325

第 1 章

初探深度学习

本章主要内容：

- 了解什么是深度学习
- 了解深度学习的加速硬件
- 了解深度学习的软件框架

我们暂且不管深度学习是什么，深度学习有多强大。作为信息行业的杰出的工程师，首先需要知道深度学习真正带来的是什么？未来，深度学习对社会以及各个行业会带来什么影响？拥有大量深度学习人才的中国企业将会在世界上充当何种角色？深度学习又会给我们带来什么样的机遇与挑战？

“路漫漫其修远兮”。深度学习需要大量的数据和庞大的计算资源，而这就将我们的视线带入 GPU 的世界。如果时至今日，你还以为英伟达（NVIDIA）只是一家卖显卡的公司，那就显得有点孤陋寡闻了。因为如今的特斯拉自动驾驶系统 Autopilot 2.0、行车预警系统、无人采集系统、智能物流系统等，到处都充斥着 NVIDIA 的身影。正是因为深度学习，NVIDIA 才会成为新一轮人工智能公司中的独角兽。

深度学习在各个领域带来的变革才刚刚开始，如黎明破晓一样让人心潮澎湃。近期关于无人驾驶、辅助驾驶、智能音箱的新闻越来越多，无论大公司，还是新创公司都在谈人工智能，为什么呢？因为深度学习极大地降低了技术的门槛。只要拥有充足的数据，加点硬件成本和时间成本，就可以利用深度学习这一技术实现各种方案，这是新创公司实现弯道超车的机会。

在本章中，我们将会了解到什么是深度学习，探索深度学习的应用，知晓深度学习的强大。近年来深度学习呈爆发式增长，主要得益于其相关硬件加速器和软件平台的迅猛发展。因此本章将会讲解深度学习最常用的硬件平台，及其对应的软件架构平台。通过对深度学习初步的了解，相信上面的问题将会给我们带来更多关于技术的思考。

1.1 什么是深度学习

2016 年年初，由 Google DeepMind 开发的 AlphaGo 在围棋大战中打败了韩国的围棋大师李世石。就如同 1997 年 IBM 的“深蓝”计算机战胜了世界首席国际象棋大师卡斯帕罗夫一样，媒体开始铺天盖地般地宣传人工智能时代的来临。同时，资本开始追捧与人工智能产业相关的公司，一时间收购并购人工智能企业的狂潮席卷而来。

在描述 DeepMind 胜利的时候，很多媒体都会把人工智能（Artificial Intelligence）、机器学习（Machine Learning）和深度学习（Deep Learning）混为一谈。虽然从严格定义上来说，DeepMind 在 AlphaGo 程序中对上述 3 种技术都有所使用，但其真正使用得更多的是深度学习。

下面来了解人工智能、机器学习、深度学习这三者之间的关系。如图 1-1 所示，人工智能包含着机器学习，而深度学习则是机器学习的重要分支之一，它们三者是包含与被包含的关系。

从 20 世纪 50 年代，人工智能第一次提出至今，人工智能的问题基本上已经定

型，大致分为6个具体的方向：问题求解、知识推理、规划问题、不确定性推理、通信感知与行动、学习问题。而机器学习主要有3个方向：分类、回归、关联性分析。最后到深度学习则是对机器学习中的神经网络进行深度拓展。

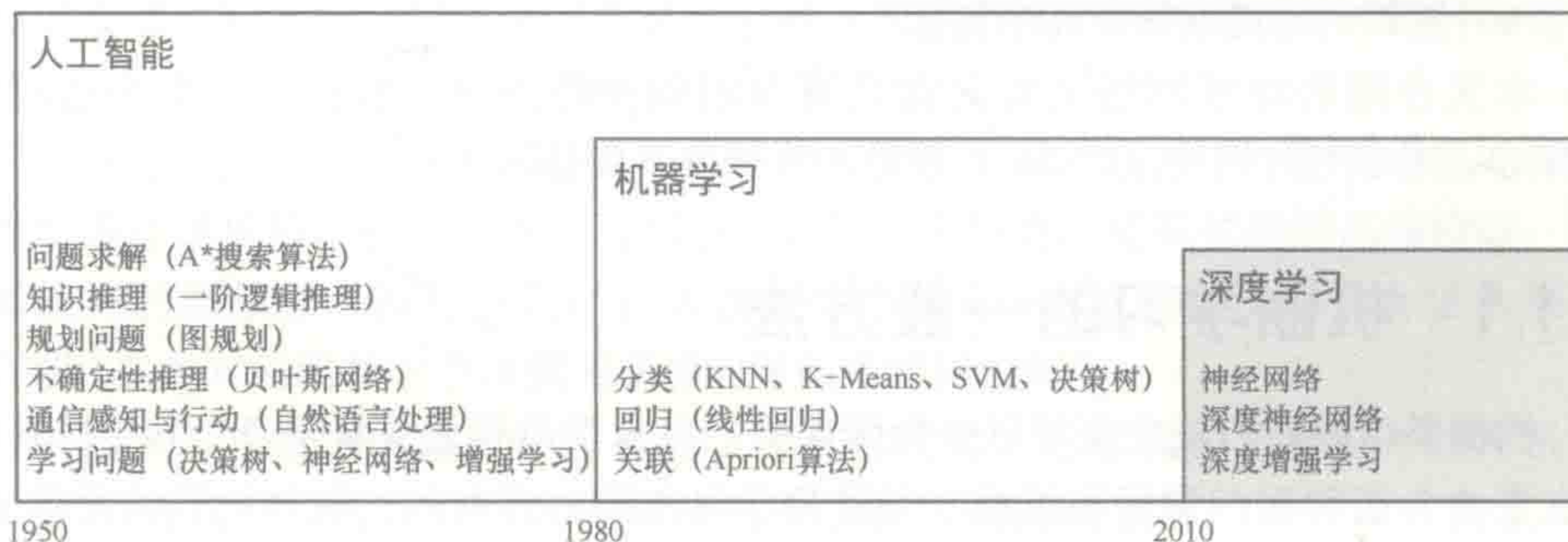


图 1-1 深度学习、机器学习、人工智能之间的关系。人工智能的概念从 1950 年左右开始被提出，而机器学习则是在 1980 年左右被提出的，最近热门的深度学习则是在 2010 年左右被提出的

人工智能实际上包含了我们经常使用到的算法，例如在问题求解方面，最经典的算法是 A* 搜索算法和 a-b 剪枝算法，又如人工智能中的学习问题则包含了机器学习的大部分内容。现阶段已经有很多资料介绍机器学习相关的算法，较为著名的机器学习十大算法有：决策树、支持向量机 SVM、随机森林算法、逻辑回归、朴素贝叶斯、KNN 算法、K-means 算法、Adaboost 算法、Apriori 算法、PageRank 算法。

深度学习的定义在过去的十年一直在不断变化，其中大部分学者认为深度学习应该“具有两层以上的神经网络”，图 1-2 所示为深度学习的神经网络示例，其特点是：

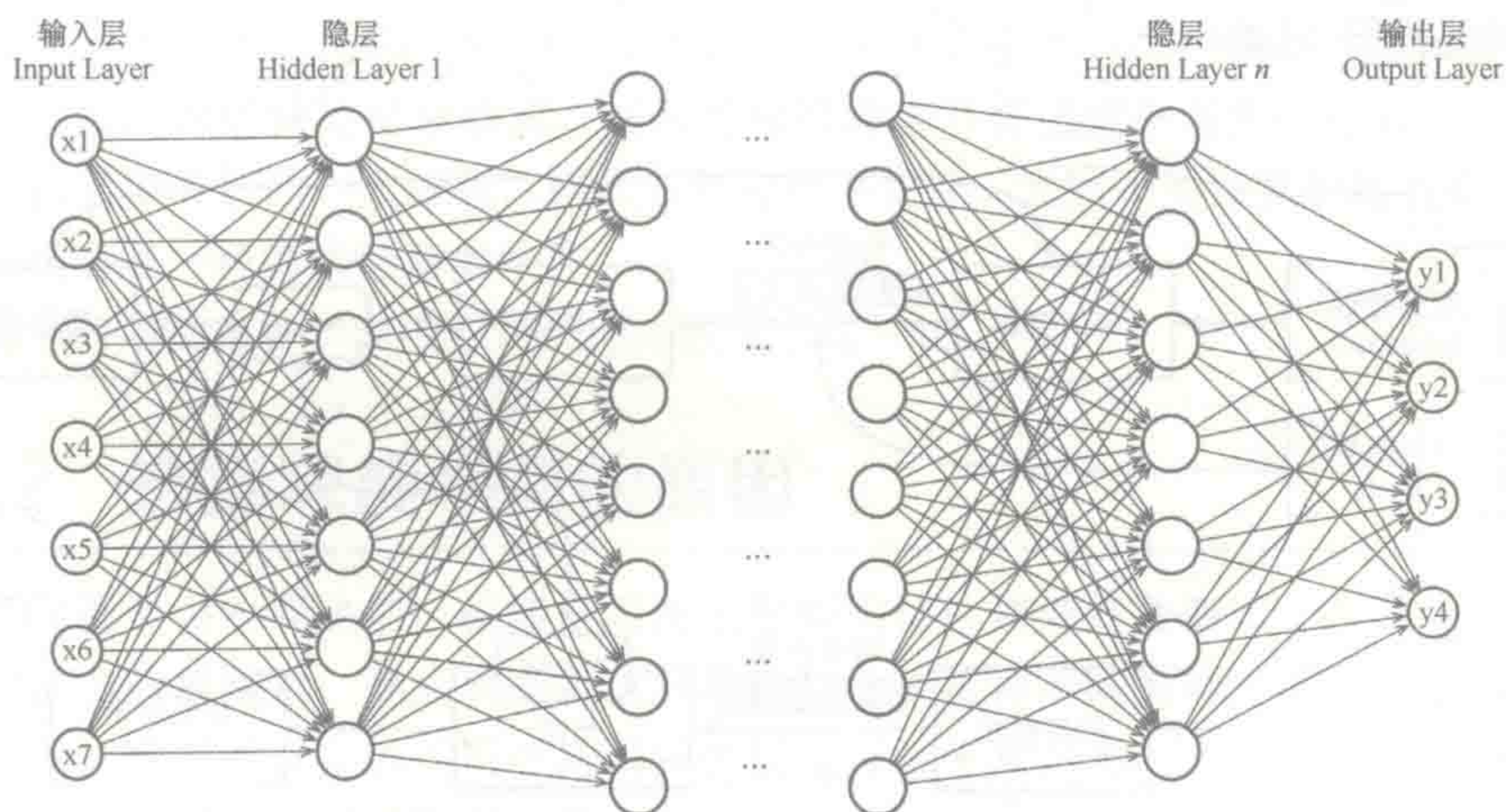


图 1-2 深度神经网络示例图。深度神经网络其实是深度学习的基础，深度学习的应用和技术绝大部分都是基于深度神经网络框架

- 具有更多的神经元;
- 具有更复杂的网络连接方式;
- 拥有惊人的计算量;
- 能够自动提取数据高维特征。

本文介绍的深度网络主要是指具有深层的神经网络，包括：人工神经网络（ANN）、卷积神经网络（CNN）、循环神经网络（RNN）。

1.1.1 机器学习的一般方法

机器学习按照方法主要可以分为两大类：监督学习和无监督学习。其中监督学习主要由分类和回归等问题组成，无监督学习主要由聚类 and 关联分析等问题组成。深度学习则属于监督学习当中的一种。

机器学习中的监督学习是指使用算法对结构化或者非结构化的有标注的数据进行解析，从数据中学习，获取数据中特定的结构模型，然后使用这些结构或者模型来对未知的新数据进行预测。通俗来说，监督学习就是通过对数据进行分析，找到数据的表达模型，有了这个模型就可以对新输入的数据套用该模型来做决策。

图 1-3 为监督学习的一般方法，主要分为训练和预测阶段。在训练阶段（对数据进行分析的阶段），首先需要根据原始的数据进行特征提取，这个过程叫作特征工程。得到特征后，我们可以使用决策树、随机森林等模型算法去分析数据之间的特征或者关系，最终得到关于输入数据的模型（Model）。在预测阶段，同样按照特征工程的方法提取了数据后，使用训练阶段得到的模型对特征向量进行预测，最终得到所属的标签（Labels）。

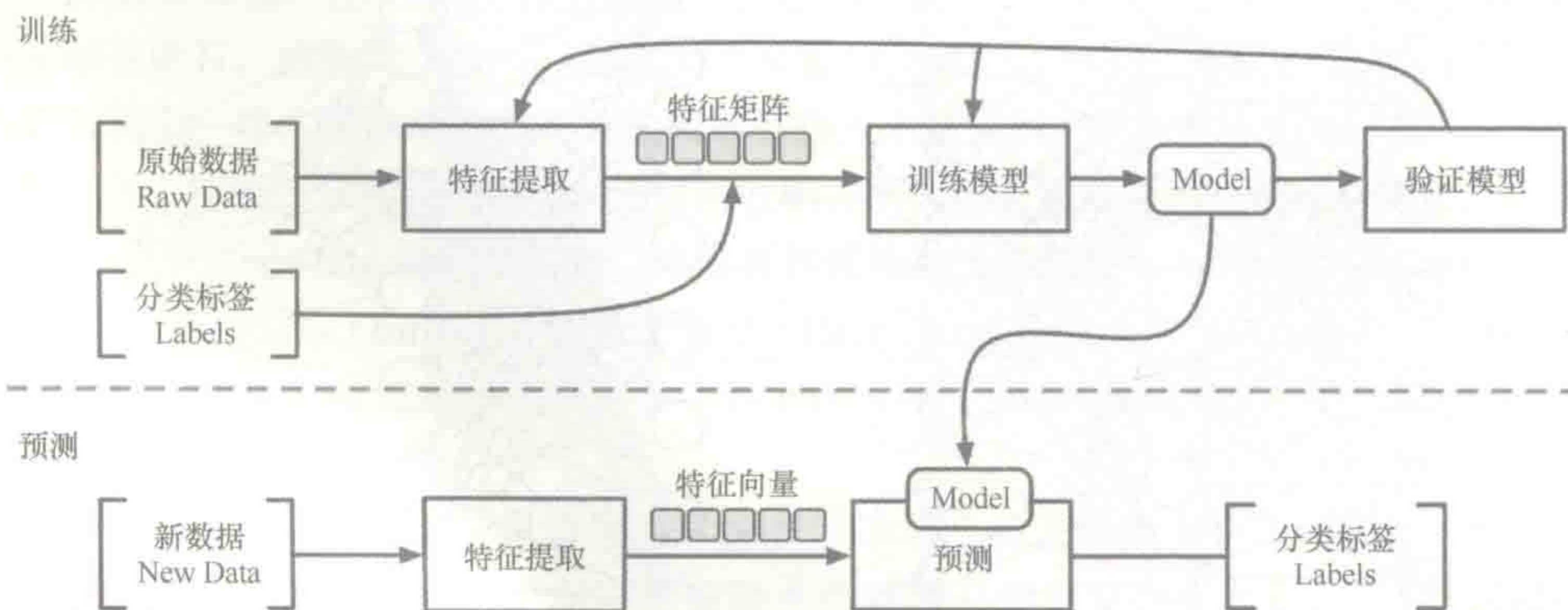


图 1-3 监督学习一般方法，分为训练阶段和预测阶段，训练阶段的目标是得到训练模型，而预测阶段的目标是使用训练模型对输入数据进行预测