

人工智能 人才培养系列

深度学习

原理与TensorFlow实践

◎ 黄理灿 编著

既有由浅入深的理论知识，又有从入门到专业的应用编程技术讲解。

内容涵盖

深度学习理论、Python 编程语言以及 TensorFlow 编程知识和代码解读。

提供 PPT 课件、源代码、在线题库等丰富配套资源。

 中国工信出版集团

 人民邮电出版社
POSTS & TELECOM PRESS

深度学习

原理与 TensorFlow 实践

RFID

© 黄理灿 编著

人民邮电出版社
北京

图书在版编目 (CIP) 数据

深度学习原理与 TensorFlow 实践 / 黄理灿编著. --
北京: 人民邮电出版社, 2019. 8
ISBN 978-7-115-50996-3

I. ①深… II. ①黄… III. ①人工智能—算法—研究
IV. ①TP18

中国版本图书馆CIP数据核字(2019)第051259号

内 容 提 要

本书首先介绍当前学术界和工业界的深度学习核心知识, 具体内容包括机器学习概论、神经网络和深度学习原理; 然后介绍深度学习的实现及深度学习框架 TensorFlow, 具体内容包括 Python 编程基础、TensorFlow 编程基础、TensorFlow 模型、TensorFlow 编程实践、TensorFlowLite 和 TensorFlow.js、TensorFlow 案例——医学应用和 Seq2Seq+attention 模型及其应用案例。

本书的最大特色是既有由浅入深的理论知识, 又有从入门到专业的技术应用。本书涵盖了深度学习的理论、Python 编程语言、TensorFlow 编程知识以及代码解读, 为深度学习初学者以及进阶人员提供了详尽的必备知识。

本书既可以作为高等院校本科高年级以及研究生的人工智能课程教材, 也可以作为应用领域技术人员、工程技术人员和科学研究工作者的参考资料。

-
- ◆ 编 著 黄理灿
责任编辑 罗 朗
责任印制 陈 犇
 - ◆ 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>
三河市中晟雅豪印务有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 22.75 2019 年 8 月第 1 版
字数: 623 千字 2019 年 8 月河北第 1 次印刷
-

定价: 69.80 元

读者服务热线: (010) 81055256 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

人工智能是当前影响人类的重大技术。近年来，由于计算机软硬件的进步，从早期的单层神经网络发展起来的多层神经网络技术（深度学习）得到了迅速的发展和广泛的应用。其应用范围包括自动驾驶汽车、语言翻译、对象识别、医疗诊断、自动聊天、自动写作、艺术品生成、植物识别等领域。TensorFlow 是开源的深度学习实现框架，具有很大的下载量和用户数，不仅可以用于研究和试验，还可以直接用于生产部署。

由于深度学习的快速发展，学术界、工业界需要大量的人才，全球人工智能领域的人才缺口也非常大。人工智能领域相关人才分为两类：研究新理论与新模型的高级研究人才和各领域的应用人才。研究新理论与新模型的高级研究人才需要深厚的理论基础，而应用人才则要利用已有的模型对数据进行训练和优化，相对而言，应用人才不需要深入的专业理论知识。由于深度学习和 TensorFlow 是最近发展的新技术，市场上还少见涵盖原理与实践的教材。我们在这里尝试编写既适合初学者又适合已迈入高级研究者门槛的大学本科以及研究生的深度学习教材。希望本书不仅适用于计算机专业人员，也适用于非计算机专业人员（如生物专业、医学专业的读者）。

本书共 11 章，包括深度学习理论和 TensorFlow 实践。读者可以根据实际情况选学带“*”号章节。

第 1 章概要性地介绍深度学习的发展历程和 TensorFlow 的应用现状。

第 2 章介绍机器学习所需的数学知识、机器学习方法以及数据预处理方法。机器学习方法包括监督学习、非监督机器学习、半监督机器学习和强化学习。

第 3 章介绍神经网络基础知识和神经网络模型。

第 4 章介绍多层感知机神经网络、卷积神经网络、循环神经网络、深度置信网络以及深度学习框架。

第 5 章介绍 Python 安装、Jupyter Notebook 器安装使用、Python 编程基础知识、Python 标准库和 Python 机器学习库。

第 6 章介绍 TensorFlow 发展历程与演进、TensorFlow 的搭建配置、TensorFlow 编程基础知识、TensorFlow 系统架构及源码结构、Eager Execution 以及简单的 TensorFlow 示例代码。

第7章介绍 TensorFlow 模型编程模式，读取数据，TensorFlow 模型搭建，TensorFlow 模型训练，TensorFlow 评估，TensorFlow 模型载入、保存以及调用，可视化评估工具 Tensorboard 以及编程示例——鸢尾花分类。

第8章介绍 TensorFlow 编程实践，包括 MNIST 手写数字识别、Fashion-MNIST 以及 RNN 简笔画识别。

第9章介绍用于移动平台的 TensorFlow Lite 和用于浏览器的 TensorFlow.js。

第10章介绍 TensorFlow 案例——医学应用，主要介绍开源医学图像分析平台 DLTK。

第11章介绍 Seq2Seq 模型、TensorFlow 自动文本摘要生成和聊天机器人示例。

为了方便读者学习和掌握所学知识，每章最后附有习题。

由于水平有限，书中难免有不妥之处，恳请读者批评指正。

黄理灿

目 录

第 1 章 绪论	1
1.1 引言.....	2
1.2 深度学习的发展历程.....	3
1.3 TensorFlow 应用现状.....	5
习题.....	6
第 2 章 机器学习概论	7
2.1 机器学习相关的数学知识.....	8
2.1.1 微积分.....	8
2.1.2 线性代数.....	11
2.1.3 概率论.....	14
2.2 机器学习方法.....	15
2.2.1 监督学习.....	16
2.2.2 无监督学习.....	24
2.2.3 半监督学习.....	26
2.2.4 强化学习.....	28
2.3 数据的预处理方法.....	31
习题.....	34
第 3 章 神经网络	35
3.1 神经网络基础知识.....	36
3.1.1 MP 模型.....	36
3.1.2 感知机.....	38
3.1.3 三层感知机.....	41
3.2 神经网络模型.....	53
3.2.1 径向基函数网络.....	54
3.2.2 Hopfield 神经网络.....	56
3.2.3 Elman 神经网络.....	56
3.2.4 玻尔兹曼机.....	57
3.2.5 自动编码器.....	60
3.2.6 生成对抗网络.....	62
习题.....	64
第 4 章 深度学习	65
4.1 多层感知机神经网络.....	66
4.2 激活函数、损失函数和过拟合.....	71
4.2.1 激活函数.....	71
4.2.2 损失函数(代价函数).....	74
4.2.3 防止过拟合.....	78
4.3 卷积神经网络.....	80
4.3.1 卷积神经网络原理.....	81
4.3.2 *卷积神经网络 BP 算法的数学 推导.....	86
4.4 循环神经网络.....	89
4.4.1 循环神经网络模型原理.....	90
4.4.2 *BPTT 算法.....	91
4.4.3 双向循环神经网络.....	95
4.4.4 深度循环神经网络.....	96
4.4.5 长短时记忆网络.....	96
4.4.6 门控循环单元网络.....	98
4.5 深度置信网络.....	99
4.5.1 RBM 原理.....	99
4.5.2 RBM 求解算法.....	100
4.5.3 对比散度算法.....	101
4.5.4 *公式推导.....	101
4.5.5 深度置信网络训练.....	105
4.6 深度学习框架.....	106
4.6.1 TensorFlow.....	107
4.6.2 Caffe.....	107
4.6.3 Theano.....	108

4.6.4 Keras	109	6.2.3 在 Linux 系统环境下安装 TensorFlow	159
习题	110	6.3 TensorFlow 编程基础知识	159
第 5 章 Python 编程基础	111	6.3.1 张量	159
5.1 Python 环境搭建	112	6.3.2 符号式编程	160
5.1.1 Python 安装	112	6.3.3 变量和常量	161
5.1.2 Jupyter Notebook 编程器安装 使用	112	6.3.4 会话	161
5.2 Python 编程基础知识	117	6.3.5 占位符、获取和馈送	162
5.2.1 Python 标识符	117	6.3.6 Variable 类	163
5.2.2 Python 标准数据类型	118	6.3.7 常量、序列以及随机值	164
5.2.3 Python 语句	118	6.3.8 执行图	166
5.2.4 Python 运算符	119	6.3.9 操作运算	167
5.2.5 代码组	121	6.3.10 基本数学函数	170
5.2.6 Python 流程控制	122	6.3.11 矩阵数学函数	171
5.2.7 Python 函数	123	6.3.12 张量数学函数	176
5.2.8 Python 模块	126	6.3.13 张量 Reduction 操作	176
5.2.9 Python 类	127	6.3.14 累加和累积	179
5.2.10 命名空间和作用域	131	6.3.15 张量拆分操作	179
5.3 Python 标准库	132	6.3.16 序列比较与索引	182
5.4 Python 机器学习库	132	6.3.17 张量数据类型转换	183
5.4.1 NumPy	132	6.3.18 TensorFlow 张量形状的确 定与改变	184
5.4.2 SciPy	140	6.4 TensorFlow 系统架构及源码 结构	185
5.4.3 pandas	143	6.5 Eager Execution	188
5.4.4 scikit-learn	148	6.6 TensorFlow 示例代码	189
习题	153	6.6.1 简单回归拟合	189
第 6 章 TensorFlow 编程 基础	155	6.6.2 波士顿房价预测	191
6.1 TensorFlow 的发展历程与演进	156	习题	193
6.2 TensorFlow 的搭建配置	158	第 7 章 TensorFlow 模型	194
6.2.1 在 Windows 系统环境下安装 TensorFlow	158	7.1 TensorFlow 模型编程模式	195
6.2.2 在 Mac OS 系统环境下安装 TensorFlow	158	7.1.1 tf.nn 模块	195
		7.1.2 tf.layers 模块	207

7.1.3	tf.estimator 模块.....	210	8.2.2	Fashion MNIST 代码.....	259
7.1.4	tf.keras 模块.....	211	8.3 RNN 简笔画识别	265	
7.2 读取数据	212		习题	275	
7.2.1	载入数据.....	212	第 9 章 TensorFlow Lite 和 TensorFlow.js	276	
7.2.2	创建迭代器.....	214	9.1 TensorFlow Lite	277	
7.2.3	使用 dataset 数据.....	216	9.1.1	转化训练好的模型为.tflite 文件.....	278
7.3 TensorFlow 模型搭建	218		9.1.2	编写自定义操作代码.....	279
7.4 TensorFlow 模型训练	220		9.1.3	在 TensorFlow Lite 的移动端进行安卓开发.....	280
7.4.1	损失函数——tf.losses 模块.....	220	9.1.4	在 TensorFlow Lite 的移动端进行 iOS 开发.....	283
7.4.2	优化器——tf.train 模块.....	220	9.2 TensorFlow.js	284	
7.4.3	训练示例.....	222	9.2.1	TensorFlow.js JavaScript 库引入.....	284
7.5 TensorFlow 评估	222		9.2.2	TensorFlow.js 基础知识.....	285
7.5.1	评价指标.....	222	9.2.3	TensorFlow.js 示例.....	289
7.5.2	评估函数——tf.metrics 模块.....	225	习题	302	
7.6 TensorFlow 模型载入、保存及调用	227		第 10 章 TensorFlow 案例——医学应用	303	
7.7 可视化分析和评估模型	229		10.1 开源医学图像分析平台 DLTK 的安装运行	304	
7.7.1	tf.summary 模块.....	229	10.2 开源医学图像分析平台 DLTK 的使用	305	
7.7.2	TensorBoard 可视化评估工具.....	229	10.3 开源医学图像分析平台 DLTK 案例	310	
7.7.3	TensorBoard 使用案例.....	230	10.4 开源医学图像分析平台 DLTK 模型	312	
7.8 示例——鸢尾花分类	239		习题	323	
习题	242				
第 8 章 TensorFlow 编程实践	243				
8.1 MNIST 手写数字识别	244				
8.1.1	使用 tf.nn 模块实现 MNIST 手写数字识别.....	245			
8.1.2	使用 tf.estimator 模块实现 MNIST 手写数字识别.....	248			
8.2 Fashion MNIST	253				
8.2.1	Keras 序列模型.....	253			

第 11 章 Seq2Seq+attention

模型及其应用案例 324

11.1 Seq2Seq 和 attention 模型 325

11.2 TensorFlow 自动文本摘要

生成 327

11.2.1 TextSum 安装运行 328

11.2.2 TextSum 整体结构 329

11.3 聊天机器人 350

11.3.1 DeepQA 350

11.3.2 Stanford TensorFlow

Chatbot 356

习题 356

01

第1章 绪论

本章概要性地介绍了深度学习的发展历程和TensorFlow的应用现状,使读者对深度学习与TensorFlow有初步的印象。本章提到的相关概念会在本书后面章节中详细介绍。

1.1 引言

人工智能是当前影响人类的重大技术。近年来，由于计算机软硬件的进步，从早期的单层人工神经网络发展起来的多层人工神经网络技术（深度学习）得到了有效的发展和广泛的应用，如自动驾驶汽车、语言翻译、对象识别、医疗诊断、自动聊天、自动写作、艺术品生成、植物识别，等等。科技界、工业界投入大量的人力和物力对深度学习进行研究并开发了许多软件工具以及专用的人工智能芯片。Google、Microsoft、Facebook、Amazon、百度、腾讯、阿里巴巴等中外公司都在积极布局，争夺深度学习的战略高点。

未来 10 年，人工智能将给人类世界带来颠覆性的变化，也将变得无处不在。人工智能技术将进入大规模的商用阶段，人工智能产品将全面进入消费级市场。机器翻译、智能音箱、面部识别、智能助手等产品将普及。基于深度学习的人工智能的认知能力将达到人类专家顾问级别。在金融投资领域，人工智能已经有取代人类专家顾问的迹象。高盛已经开始布局智能投资顾问，苏格兰皇家银行也宣布用投资顾问取代 500 名传统理财师的工作。与保险业相结合，人工智能将为保险机构提供顾问服务。与医疗领域相结合，人工智能将为医生、病人提供咨询服务。人工智能与不同产业的结合正让人工智能逐步成为一种可以购买的商品。人工智能技术将严重地冲击劳动密集型产业，改变全球经济生态。

人工智能是如何获得“智能”的呢？这主要归功于一种实现人工智能的方法——机器学习。机器学习最基本的做法是使用算法来解析数据，并从数据中学习，然后对真实世界中的事件做出决策和预测。

传统的机器学习算法包括决策树、聚类、贝叶斯分类、支持向量机，等等。从学习方法上来分，机器学习方法可以分为监督学习、无监督学习、半监督学习和强化学习。

机器学习的一个分支是借鉴生物大脑神经元系统的人工神经网络。人工神经网络通过从样本数据中学习并调整权值和偏置，从而预测真实的数据。学习权值和偏置的方法得益于反向传播算法。近年来计算机处理能力和理论的发展，将人工神经网络的层次推向了多层，有的甚至达到 1000 层以上。这种多层人工神经网络的机器学习方法被称为深度学习。人工智能、机器学习、人工神经网络和深度学习的关系，如图 1.1 所示。人工智能涵盖了一切“人工的智能技术”，而机器学习则是通过数据学习得到的“人工智能”，人工神经网络则是利用不同的人工神经网络结构并训练优化人工神经网络结构内的参数得到的“人工智能”，深度学习是指具有多层人工神经网络结构的人工神经网络学习方法。

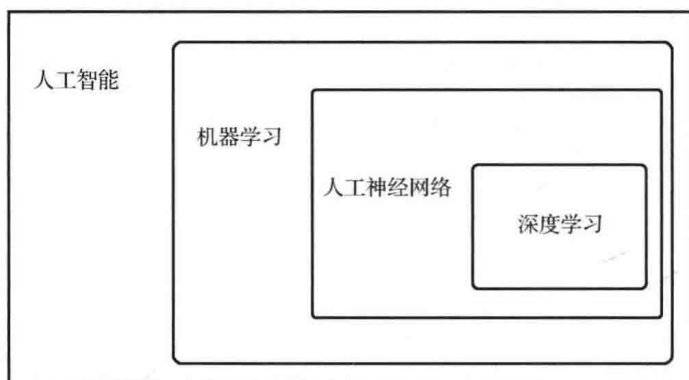


图 1.1 人工智能、机器学习、人工神经网络和深度学习的关系

深度学习需要非常强的计算能力。为了提高硬件的计算能力，Google 开发了专用的芯片 TPU (Tensor Processing Unit); NVIDIA 公司也推出了具有更强大计算能力的图形处理芯片 GPU (Graphics Processing Unit)。

同时，科技界和工业界纷纷推出了便于使用的深度学习开发工具，如 Google 的 TensorFlow、Facebook 的 Caffe 等工具。

人工智能不仅是各大公司的战略重点，也是全球各国的战略重点。未来国家之间的竞争大部分将取决于人工智能的竞争，归根到底，是人工智能的人才竞争。由于深度学习技术急剧发展，在学术界、工业界都需要大量的人才。我国深度学习的人才缺口非常大，预计将超过 500 万人。国内深度学习人才的供求比例仅为 1 : 10, 供求严重失衡。很多行业会用高薪待遇吸引深度学习的顶级人才。深度学习的人才既来源于高等院校的培养，也来自其他行业的人才转型。

1.2 深度学习的发展历程

神经网络作为当前最有前途的人工智能的一部分，经历了三个不同的发展阶段，有高潮也有低谷。

如图 1.2 所示，1943 年到 1969 年是提出神经元数学模型到单层感知机的第一阶段；经过 1969 年到 1986 年的停滞，1986 年到 1998 年是提出并应用反向传播算法的第二阶段；又经过 1998 年到 2006 年的停滞，从 2006 年开始到现在是提出深度学习并广泛应用的第三阶段。

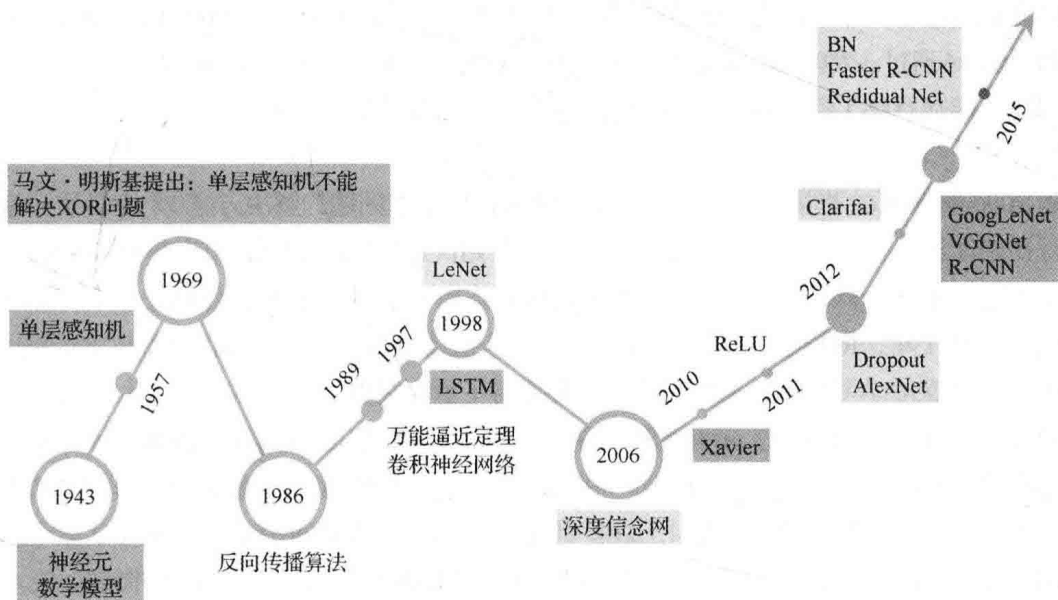


图 1.2 神经网络的发展历程

生物具有学习能力是由于生物大脑具有很多神经元。美国心理学家麦卡洛克 (McCulloch W. S.) 和数学家皮特斯 (Puts W.) 等在 1943 年参考了生物神经元的结构，发表了抽象的神经元模型——麦卡洛克-皮特斯模型 (McCulloch-Pitts Model)，简称 MP 模型，开启了人工神经网络的研究。MP 模型将神经元简化为了三个过程：输入信号线性加权、求和以及非线性激活（阈值法）。在此基础上，弗兰克·罗森布拉特 (Frank Rosenblatt) 于 1957 年提出了感知机 (Perception) 算法。该算法使用 MP 模型对输入的多维数据进行二分类，且能够使用梯度下降法从训练样本中自动学习更新权值。

1969年，马文·明斯基（Marvin Minsky）证明了感知机本质上是一种线性模型，只能处理线性分类问题，就连最简单的 XOR（异或）问题都无法正确分类。

在研究停顿了近20年以后，杰弗里·辛顿（Geoffrey Hinton）于1986年提出了适用于多层感知机（Multi-Layer Perception, MLP）的反向传播算法（BP算法），并采用 Sigmoid 函数（Sigmoid 函数是 S 型函数，常被用作神经网络的阈值函数，将变量映射到 0 和 1 之间）进行非线性映射，有效地解决了非线性分类和学习的问题。乔治·塞本柯（Geogre Cybenko, 1989年）和库尔特·霍尼克（Kurt Hornik, 1991年）证明了多层感知机（MLP）的万能逼近定理，即对于任何闭区间内的一个连续函数 f ，都可以用含有一个隐含层的 BP 网络来逼近。同年，延恩·勒昆（Yann LeCun）发明了卷积神经网络——LeNet，并将其用于数字识别。1991年，BP 算法被指出存在梯度消失的问题，即在误差梯度反向传递的过程中，后层梯度以乘法方式叠加到前层，由于 Sigmoid 函数的饱和特性，后层梯度本来就小，误差梯度传到前层时几乎为 0，因此无法对前层进行有效的学习。1997年，长短期记忆网络（LSTM）模型被发明，尽管该模型在序列建模上的特性非常突出，但由于当时正处于神经网络的下坡期，因此没有引起足够的重视。

在神经网络发展受阻的同时，其他机器学习方法得到了快速的发展。1986年，决策树方法被提出，很快 ID3、ID4、CART 等改进的决策树方法相继出现，到目前决策树仍然是非常常用的一种机器学习方法。1995年，线性支持向量机（SVM）被统计学家万普尼克（Vapnik）提出。该方法的特点有两个：由非常完美的数学理论推导而来（统计学与凸优化等）和符合人的直观感受（最大间隔）。不过，最重要的还是该方法在线性分类的问题上取得了当时最好的成绩。2000年，Kernel SVM 被提出，核化的 SVM 通过一种巧妙的方式将原空间线性不可分的问题，通过 Kernel 映射成高维空间的线性可分问题，成功解决了非线性分类的问题，且分类效果非常好。至此也进一步终结了神经网络的研究热情。

直到2006年 Hinton 提出了深层神经网络训练中梯度消失问题的解决方案（无监督预训练方法对权值进行初始化加上有监督训练微调），神经网络才开始快速发展，从此称多层神经网络为深度学习。其主要思想是先通过自学习的方法学习到训练数据的结构（自动编码器），然后在该结构上进行有监督训练微调。2011年，ReLU 激活函数被提出，该激活函数能够有效地抑制梯度消失的问题。同年，Microsoft 首次将深度学习应用在语音识别上，并取得了重大突破。

2012年，Hinton 课题组为了证明深度学习的潜力，首次参加了 ImageNet 图像识别比赛，通过其构建的卷积神经网络（CNN），AlexNet 一举夺得冠军，并在分类性能上远超第二名（采用 SVM 方法）。也正是由于该比赛，卷积神经网络吸引了众多研究者的注意。AlexNet 首次采用 ReLU 激活函数，极大地增加了收敛速度且从根本上解决了梯度消失问题；AlexNet 抛弃了“预训练+微调”的方法，完全采用有监督训练；用添加 Dropout 层方法来减小过拟合，添加局部响应归一化层（全称为 Local Response Normalization, LRN）增强泛化能力/减小过拟合，并首次采用 GPU 对计算进行加速。2015年，Hinton, LeCun, 本吉奥（Bengio）说明了（无严格论证）损失（Loss）的局部极值问题对于深层网络来说可以忽略。该论断也消除了笼罩在神经网络上的局部极值问题的阴霾。具体原因是深层网络虽然局部极值非常多，但是通过深度学习的批量梯度下降（Batch Gradient Descent）优化方法很难陷进去，而且就算陷进去，其局部极小值点与全局极小值点也非常接近，但是浅层网络却不然，其拥有较少的局部极小值点，但是很容易陷进去，且这些局部极小值点与全局极小值点相差较大。2015年，MSRA（微软亚洲研究院）的何凯明（Kaiming He）等提出了深度残差网络（Deep Residual Net），

极大地增强了深度学习网络的表达能力，并且能够轻松训练高达 150 层的神经网络。2014 年，左景贤（Kyunghyun Cho）等人提出了将循环神经网络编码器-解码器（RNN Encoder-Decoder）用于机器翻译。至此，深度学习成为学术界、工业界的热点。而深度卷积神经网络和循环神经网络也成为自动驾驶汽车、语言翻译、对象识别、医疗诊断、自动聊天、自动写作、艺术品生成、植物识别等应用的主流神经网络。

1.3 TensorFlow 应用现状

TensorFlow 是 Google 公司开发并开源的机器学习框架。从 2010 年开始，Google Brain 将 DistBelief 作为第一代专有的机器学习系统。50 多个团队在 Google 和其他 Alphabet 子公司的商业产品中部署了基于 DistBelief 的深度学习神经网络，包括 Google 搜索、Google 语音搜索、广告、Google 相册、Google 地图、Google 街景、Google 翻译和 YouTube。在此基础上，Google 在 Geoffrey Hinton 和杰夫·迪恩（Jeff Dean）的领导下，简化和重构了 DistBelief 的代码库，使其变成一个更快、更健壮的应用级别代码库，并于 2015 年 11 月 9 日以 TensorFlow 这个名称在 Apache 2.0 开源许可证下发布。2016 年 4 月 14 日，Google 发布了重大的更新版本——分布式 TensorFlow 0.8。同年 6 月，发布了 TensorFlow 0.9 版本，增加了对 iOS 的支持。

2017 年 2 月 15 日晚，Google 举办了第一届 TensorFlow 开发者峰会。在该峰会上，Google 正式发布了 TensorFlow 1.0 版本。此后 TensorFlow 版本更新很快，目前已发布到 TensorFlow 1.9 版本。

TensorFlow 支持高性能处理器，包括各种 CPU、GPU 和 Google 的 TPU。TensorFlow 支持分布式运行。TensorFlow 1.0 版本在性能上有了很大的提升，可以做到在 64 块 GPU 上运行时达到 58 倍的加速。

TensorFlow 支持各种设备，能够很好地在服务器、桌面 PC、移动端（如安卓设备、iOS 等）以及云平台等设备上运行。

TensorFlow 支持多种不同的编程语言接口，如 Python、C++、Go、Java 等。

TensorFlow 包括像 TensorBoard 这类的工具，能够有效地提高深度学习网络开发的效率。

TensorFlow 封装了高级 API，在灵活性、可扩展性、可维护性上具有很大的优势。例如，Layers API 封装了一些层的操作；Estimator API 封装了一些更高层的操作（包括 train 和 evaluate 操作），用户可以通过几行代码快速构建训练和评估过程。

Keras 是一个可以在很多平台上应用的深度学习框架，深受广大研究人员和工作者的喜爱。TensorFlow 集成了 Keras，使得 TensorFlow 很容易调用 Keras 代码。

TensorFlow 目前已经成为在机器学习、深度学习项目中最受欢迎的框架之一。很多公司和产品都使用 TensorFlow。

TensorFlow 是使用计算图方式进行计算的框架。计算图中的节点代表数学运算，而计算图中的边则代表在这些节点之间传递的多维数组（张量）。计算图具有高性能，但编程较难的特点。TensorFlow 也发展了将用传统编程方法编写的代码直接转化为图计算的工具。TensorFlow 不仅用于科学研究，也可直接用于工业生产。

Google 公司本身很多的应用都采用了 TensorFlow 框架，例如 Gmail、Google Play Recommendation、Search、Translate、Map 等。在医疗方面，使用 TensorFlow 搭建了根据视网膜来预防因糖尿病致盲的

系统；在音乐、绘画领域使用 TensorFlow 构建了深度学习模型来帮助人类更好地理解艺术；使用 TensorFlow 框架构建的自动化的海洋生物检测系统，可以帮助科学家了解海洋生物的情况；在移动设备上使用 TensorFlow 做翻译工作，等等。以下是几个利用 TensorFlow 做应用的具体例子。

- **Magenta:** 利用深度学习做与艺术相关的工作。
- **AlphaGo:** 在围棋比赛上打败人类，然后升级版的 Master 保持了围棋比赛中连续 60 盘不败的记录。
- **WaveNet:** 语音音频合成。可以利用深度学习将文本用合成语音“说”出来。
- 天体物理学家使用 TensorFlow 分析开普勒任务中的大量数据，以期发现新的行星。
- 医学研究人员使用 TensorFlow 机器学习技术评估一个人心脏病发作和中风发作的概率。
- 空中交通管制员使用 TensorFlow 来预测飞机最有可能行经的路线，以确保飞机安全着陆。
- 工程师使用 TensorFlow 分析热带雨林中的声音数据，用来检测伐木车伐木和其他非法活动。
- 科学家在非洲使用 TensorFlow 检测木薯植物疾病，从而提高木薯产量。

习题

1. 详述人工智能、机器学习、人工神经网络和深度学习之间的关系。
2. 详述深度学习的发展历程。
3. 详述深度学习在社会、经济以及科学技术领域中的应用。

02

第2章 机器学习概论

本章概要性地介绍了机器学习相关的数学基础知识、机器学习方法以及数据的预处理方法。

2.1 机器学习相关的数学知识

深度学习不同于传统的计算机系统。传统的计算机系统主要是由程序决定计算的顺序和结果；而深度学习则是利用数据进行训练，改变神经网络的权值和结构，从而得出不同的计算结果。对于研究新的神经网络模型或结构，计算机专业知识和数学知识是必不可少的；然而对具体领域应用来说，领域知识也许更重要，而计算机专业知识和数学知识不是必需的。当然，具有相关的数学和计算机领域的知识对深度学习的理解和实践应用具有帮助。这里，将对机器学习相关的数学知识进行简单的介绍。

2.1.1 微积分

对于深度学习，经常需要训练数据从而使标记值与预期值的误差尽量小，如反向传播算法。理解反向传播算法需要一些微积分知识。

导数 (Derivative) 是微积分中的重要基础概念。当函数 $y=f(x)$ 的自变量 x 在一点 x_0 上产生一个增量 Δx 时，函数输出值的增量 Δy 与自变量增量 Δx 的比值在 Δx 趋于 0 时的极限值 a 如果存在， a 即为在 x_0 处的导数，记作 $f'(x_0)$ 或 $df(x_0)/dx$ 。

$$f'(x_0) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

或者

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$$

导数是函数的局部性质。一个函数在某一点的导数描述了该函数在这一点附近的变化率。如果函数的自变量和取值都是实数的话，函数在某一点的导数就是该函数所代表的曲线在这一点上的切线斜率，导数的几何意义如图 2.1 所示。

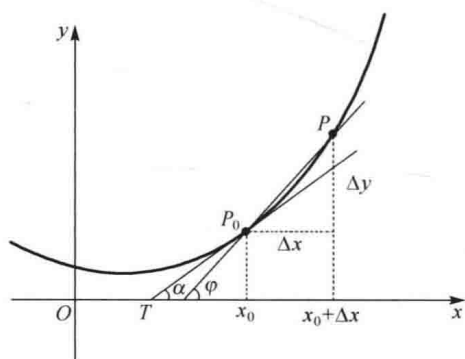


图 2.1 导数的几何意义

1. 导数公式

以下列出了一些常用的导数公式。若需了解公式证明，可参考高等数学教科书。

$$y = C, y' = 0$$

$$y = x^n, y' = nx^{n-1}$$

$$y = \sin x, y' = \cos x$$

$$y = \cos x, y' = -\sin x$$

$$y = \tan x, y' = \frac{1}{\cos^2 x} = \sec^2 x$$

$$y = \cot x, y' = -\frac{1}{\sin^2 x} = -\csc^2 x$$

$$y = \sec x, y' = \sec x \cdot \tan x$$

$$y = \csc x, y' = -\csc x \cdot \cot x$$