



**华育兴业**

战略性新兴产业人才教育解决方案专家

教育部-华育兴业

产学研合作协同育人项目成果

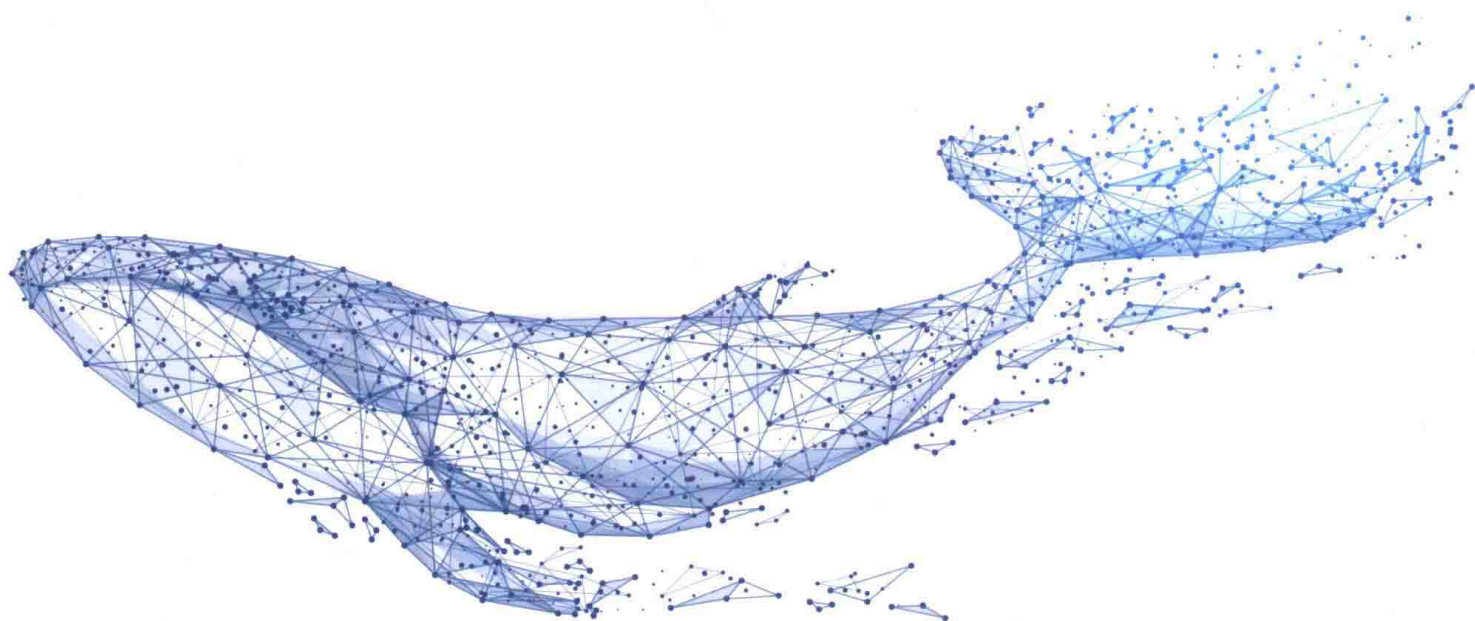
人工智能技术丛书

# A First Course in Deep Learning

# 深度学习基础教程

赵宏 ◎ 主编

于刚 吴美学 张浩然 屈芳瑜 王鹏 ◎ 参编



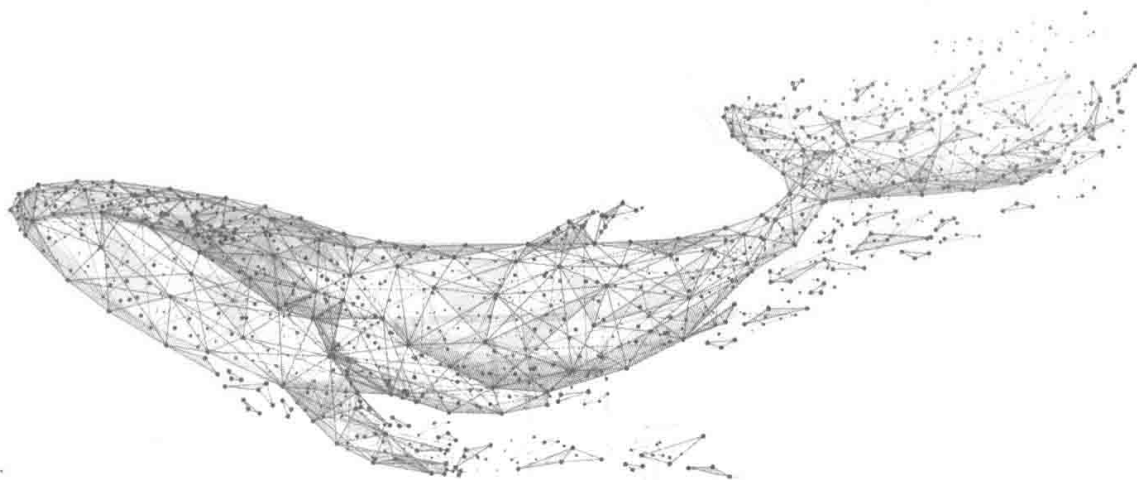
机械工业出版社  
China Machine Press

# A First Course in Deep Learning

# 深度学习基础教程

赵宏 © 主编

于刚 吴美学 张浩然 屈芳瑜 王鹏 © 参编



机械工业出版社  
China Machine Press

## 图书在版编目 (CIP) 数据

---

深度学习基础教程 / 赵宏主编. -- 北京: 机械工业出版社, 2021.7

(人工智能技术丛书)

ISBN 978-7-111-68732-0

I. ①深… II. ①赵… III. ①机器学习 - 教材 IV. ①TP181

中国版本图书馆 CIP 数据核字 (2021) 第 140992 号

---

深度学习是当前人工智能领域的热点。本书根据高等院校理工科专业学生的学习需求, 介绍深度学习的相关概念, 培养学生利用基于各类深度学习架构的人工智能算法来分析和解决相关专业问题的能力。本书内容包括深度学习概述、人工神经网络基础、卷积神经网络和循环神经网络、生成对抗网络和深度强化学习、计算机视觉以及自然语言处理。

本书可作为高等院校理工科相关专业深度学习、人工智能等课程的教材, 也可作为技术人员的参考书或自学读物。

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 朱 劼

责任校对: 马荣敏

印 刷: 三河市宏达印刷有限公司

版 次: 2021 年 8 月第 1 版第 1 次印刷

开 本: 186mm×240mm 1/16

印 张: 12

书 号: ISBN 978-7-111-68732-0

定 价: 59.00 元

---

客服电话: (010) 88361066 88379833 68326294

投稿热线: (010) 88379604

华章网站: [www.hzbook.com](http://www.hzbook.com)

读者信箱: [hzjsj@hzbook.com](mailto:hzjsj@hzbook.com)

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东



### 赵宏

博士，南开大学计算机学院教授、硕士生导师。校级教学团队带头人，2013–2017 年教育部高等学校文科计算机基础教学指导分委员会委

员，主要负责大学计算机通识课程的教学与研究。主持 23 项教改项目，主编教材 10 本。获得中国信息协会“2019–2020 年度在线教育发展贡献人物”、教育部在线教育研究中心“智慧教学之星”、南开大学教学名师、南开大学“课程思政”优秀典型课程、校级教学成果奖等多项奖励和荣誉称号。主要进行空气质量数值预报模型方面的应用研究，承担科研项目 30 余项，发表教学和科研论文 50 余篇。



### 于刚

博士，南开大学计算机学院讲师。研究方向包括网络安全与数据分析、深度学习、人工智能等，近年来在深度学习和人工智能领域参与了国

家级、省部级产学研合作项目 6 项，参与编写教材 5 部，在国内外知名期刊和会议上发表学术及教学论文 10 余篇。



### 吴美学

南开大学计算机学院硕士研究生，主要从事数据挖掘的研究与应用。参与大学生成绩预测，并在国际计算机科学与教育信息化会议（IEEE

ICCSE）发表相关论文；参与旅游推荐系统的研究，通过构建旅游知识图谱提高旅游推荐模型的准确率。对深度学习、知识图谱有一定的研究和应用基础。

## 作者介绍



### 张浩然

南开大学网络空间安全学院硕士研究生，主要从事目标检测方向的研究。曾参与“火灾现场数据处理关键技术研究”项目，并在国际计算机科学与教育信息化会议（IEEE ICCSE）发表相关论文。对计算机视觉领域的深度神经网络有一定的研究和应用基础。



### 屈芳瑜

南开大学计算机学院硕士研究生，主要从事大数据技术的研究与应用。参与大学生体质健康水平与学业成绩相关关系的研究，并在国际计算机科学与教育信息化会议（IEEE ICCSE）发表相关论文；参与完成“沈阳市大气污染源 GIS 表征系统”开发项目。对机器学习、神经网络有一定的研究和应用基础。

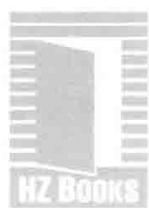


### 王鹏

北京华育兴业科技有限公司首席技术官。主要从事金融、电信、电力及教育等行业的海量数据分析和挖掘工作。中国大数据与智能计算产业联盟副秘书长，甲骨文特聘专家，黑龙江省人工智能学会理事等。曾就职于中国普天信息技术研究院，担任架构师、技术总监等职务。代表企业主导网络存储国家行业标准制定，拥有专利 2 项。



关注华育兴业公众号  
获取更多实验资源



## 华章图书

一本打开的书，一扇开启的门，  
通向科学殿堂的阶梯，托起一流人才的基石。

[www.hzbook.com](http://www.hzbook.com)

P R E F A C E

# 前 言

随着互联网与大数据时代的来临，人工智能技术已经成为当前科技发展的重要组成部分，并在工业智能制造、农业自动化生产以及商业贸易与金融宏观调控等领域得到广泛应用。基于人工智能技术的各类产品已经渗透到我们的生产、生活、教育、娱乐、医疗等各个方面，并逐渐改变着人们的思维习惯和行为方式。可以说，人工智能技术的发展与推广对人类社会的影响之深远是前所未有的。

在当前的人工智能研究领域中，深度学习（Deep Learning）是一项备受关注的关键技术，并已成为热点话题。深度学习属于机器学习范畴，是构成复杂人工神经网络结构的要素。从单纯的技术角度来看，深度学习中的“深度”是指所配置的人工神经网络模型中的隐含层在纵向有较深的结构。研究表明，人工神经网络的深度直接决定了它刻画现实问题的能力，因为只有复杂的深层神经元组合才能够拟合出较为复杂的函数处理过程，从而得到有效的智能输出结果。当然，随着时间的推移和人工智能技术的发展，深度学习也将被赋予更广泛、更丰富的内涵。

深度学习不是一个孤立的观念，它与人工神经网络技术息息相关。因此，要深入了解深度学习的原理与运行机制，需要先掌握人工神经网络的基本概念和内部结构等知识。卷积神经网络和循环神经网络作为深度学习中的两大类基础深度神经网络架构，就是从传统的人工神经网络结构衍生而来的。深度学习在各个关键领域的应用，如智能决策、计算机视觉以及自然语言处理等方面，是在这些基础深度神经网络架构之上，针对不同应用场景的数据规律与问题特征，做了进一步的网络架构整合与创新。

深度学习是一个比较新的概念。随着新技术的不断涌现，深度学习涉及的很多

概念的内涵和外延也在不断突破和发展，其中的一些新观点、新见解尚未在学术界及产业界形成统一的认识，相关的理论体系还不是很完善，仍然在持续创新和演变过程中。然而，深度学习又是当前人工智能领域的关键技术，也是开发各类人工智能算法的核心要素。本书力图通过阐述深度学习所涉及的基本概念、运行原理与应用分类，使读者形成较为完整的关于深度学习相关概念的认知，从而为进一步研究和开发各专业领域的人工智能算法打下良好的基础。

本书根据高等院校理工科专业特别是新工科相关专业学生的学习需求，介绍深度学习的相关概念，培养学生利用基于各类深度学习架构的人工智能算法来分析和解决相关专业问题的能力。

全书共分6章，主要内容如下。

第1章“深度学习概述”主要介绍深度学习产生的历史背景及发展历程，并简述深度学习的基本概念及应用场景。

第2章“人工神经网络基础”介绍人工神经网络的理论抽象模型，并给出神经网络中的前向传播机制和反向传播机制以及基于反向传播算法的神经网络设计流程。

第3章“卷积神经网络和循环神经网络”介绍卷积神经网络和循环神经网络这两种深度神经网络架构的基本概念、内部结构、工作机制与主要应用场景。

第4章“生成对抗网络和深度强化学习”介绍生成对抗网络的博弈与训练过程以及强化学习的基本原理和Q-Learning算法，并给出与之相关的网络模型和框架。

第5章“计算机视觉”介绍计算机视觉的图像预处理、相关算法分析和特征提取以及目标匹配等原理，还给出图像与视频分类以及目标检测常用的深度学习网络结构。

第6章“自然语言处理”介绍自然语言处理中的词嵌入算法和注意力机制，阐述文本分类、自动文本摘要和自动问答的相关技术，并给出自然语言处理在应用领域的主要模型。

为了便于读者进一步理解和掌握深度学习的基本概念和应用方法，本书每章都配备了多种类型的习题和丰富的案例，读者可以通过练习和实践进一步掌握深度学习的相关知识，提升应用能力。

本书是教育部—华育兴业产学合作协同育人项目（2019年第2批）的成果，由南开大学计算机学院公共计算机基础教学部的老师和研究生结合多年的教学、项目实践经验以及当前理工科专业学生对深度学习的学习需求编写而成。在本书的编写过程中，得到了华育兴业公司和机械工业出版社华章分社的大力支持，在此表示真诚的感谢！

本书在筹备、编写过程中参考了国内外深度学习及人工智能领域的一些开放课程网站、书籍、论坛、博客、论文和开源资料等，在此一并向作者们表示感谢。由于编者能力和时间的限制，书中难免有不妥或错误之处，恳请同行和读者斧正，在此表示真诚的谢意！

作者

2021年4月于南开园

## 目 录

前言	课后习题	21
	参考文献	22
<b>第 1 章 深度学习概述</b>	<b>第 2 章 人工神经网络基础</b>	24
1.1 深度学习的发展历程	2.1 人工神经网络的生物学基础	24
1.1.1 深度学习的历史	2.1.1 神经元的基本模型	24
1.1.2 深度学习领域的重要人物	2.1.2 突触的结构	26
1.2 深度学习的关键技术	2.2 人工神经元模型	26
1.2.1 深度学习的机理	2.2.1 人工神经元的数学模型	26
1.2.2 深度学习的三要素	2.2.2 常见的人工神经元模型	30
1.2.3 数据的特征	2.3 人工神经网络模型	34
1.2.4 深度学习的主要模型	2.3.1 神经网络的基本结构	34
1.2.5 深度学习模型的训练过程	2.3.2 神经网络的分类	36
1.2.6 深度学习模型的学习方式	2.4 神经网络的前向传播机制	39
1.2.7 深度学习的常用框架	2.5 神经网络的反向传播机制	40
1.3 深度学习网络的发展脉络及应用领域	2.6 基于反向传播算法的神经网络设计流程	43
1.3.1 深度学习网络的发展脉络	2.7 人工神经网络的参数优化问题	45
1.3.2 深度学习的应用领域	2.7.1 神经网络层数的优化问题	45

2.7.2	归一化指数函数 softmax	47	4.1.2	生成对抗网络的 基本原理	94
2.7.3	学习率	49	4.1.3	几种改进的生成 对抗网络模型	99
2.7.4	欠拟合和过拟合问题	50	4.1.4	生成对抗网络的应用	103
课后习题		52	4.2	强化学习	106
参考文献		53	4.2.1	强化学习概述	106
<b>第3章 卷积神经网络和 循环神经网络</b>		54	4.2.2	强化学习的决策过程	108
3.1	卷积神经网络	54	4.2.3	Q-Learning 算法	111
3.1.1	卷积神经网络的 基本概念	54	4.2.4	深度强化学习	112
3.1.2	卷积神经网络的结构	58	课后习题		118
3.1.3	卷积神经网络的 常用架构	65	参考文献		119
3.2	循环神经网络	72	<b>第5章 计算机视觉</b>		121
3.2.1	循环神经网络的 基本概念	72	5.1	计算机视觉概述	121
3.2.2	循环神经网络的 应用——语言模型	77	5.1.1	计算机视觉的历史	122
3.2.3	循环神经网络的梯度 问题及解决方法	80	5.1.2	计算机视觉的挑战与 机遇	123
3.2.4	循环神经网络的改进	84	5.1.3	计算机视觉常见的 数据集	125
课后习题		87	5.1.4	计算机视觉处理的 基本流程	130
参考文献		89	5.2	图像预处理	131
<b>第4章 生成对抗网络和 深度强化学习</b>		92	5.2.1	图像去噪	131
4.1	生成对抗网络	92	5.2.2	图像归一化	133
4.1.1	生成对抗网络概述	92	5.2.3	图像分割技术	134
			5.3	计算机视觉常用的网络结构	136
			5.3.1	图像分类常用的深度 学习网络结构	136

5.3.2 视频分类常用的深度 学习网络结构 .....	140	6.1.3 基础技术 .....	160
5.3.3 目标检测常用的深度 学习网络结构 .....	144	6.1.4 词嵌入算法 .....	162
课后习题 .....	152	6.1.5 N-gram 语言模型 .....	166
参考文献 .....	154	6.1.6 注意力机制 .....	167
<b>第 6 章 自然语言处理</b> .....	<b>156</b>	<b>6.2 自然语言处理的应用模型</b> .....	<b>171</b>
6.1 自然语言处理概述 .....	156	6.2.1 文本分类 .....	171
6.1.1 发展历史 .....	157	6.2.2 自动文本摘要 .....	175
6.1.2 自然语言处理的过程 .....	158	6.2.3 自动问答系统 .....	178
		6.2.4 触发字检测 .....	181
		课后习题 .....	182
		参考文献 .....	183

# 深度学习概述



2006年, Geoffrey Hinton 和 Ruslan Salakhutdinov 提出了深层神经网络训练中梯度消失问题的解决方案, 并给出了一种基于逐层贪心算法的深度信念网络模型, 实现了传统的人工神经网络向深度学习的转变, 关于神经网络的研究再次兴起。深度学习受到学术界和工业界越来越多的关注, 这进一步推动了它的研究和发展, 使其在搜索技术、数据挖掘、机器学习、机器翻译、自然语言处理、多媒体学习、语音、推荐和个性化技术以及相关领域都取得了很多成果, 深度学习在当前社会的各个领域已经有了较为广泛的实际应用。

本章主要介绍深度学习产生的历史背景及其发展历程, 并简述深度学习的基本概念、研究脉络及其在各领域中的应用。

## 1.1 深度学习的发展历程

### 1.1.1 深度学习的历史

1890年, 美国心理学家威廉·詹姆斯 (William James) 出版了《心理学原理》一书。该书详细地描述了人脑的结构及功能, 为人们研究人工神经网络提供了生物学基础。1943年, 心理学家沃伦·麦卡洛克 (Warren McCulloch) 和数学家沃尔

特·皮茨 (Walter Pitts) 参考生物神经元的结构, 发表了论文 “A logical calculus of the ideas immanent in nervous activity”, 并在该论文中提出了抽象的神经元模型 M-P。图 1.1 分别为 Warren McCulloch 和 Walter Pitts。1949 年, 唐纳德·奥尔丁·赫布 (D. Olding Hebb) 在此基础上提出了神经元的抽象数学模型。

1950 年, 艾伦·图灵 (Alan Turing) 发表论文 “计算机器和智能”, 第一次提出了机器是否可以思考的问题。他提出的图灵测试也成为人工智能哲学的重要组成部分, 即如何定义机器的智能、意识和能力等基础性问题。1956 年, 在美国达特茅斯大学召开的学术研讨会上, 科学家们讨论了关于机器智能的问题, 并正式采用了 “人工智能 (Artificial Intelligence, AI)” 这一术语。所谓人工智能, 就是由人类设计的具有一定的 “思考” 和 “意识” 能力并应用于机器的算法。人工智能概念一经提出, 就在全世界范围内掀起了研究热潮。机器学习 (Machine Learning, ML) 就是研究怎样使用计算机模拟或实现人类学习活动的科学, 它是人工智能中最具智能特征、最前沿的研究领域之一。



图 1.1 Warren McCulloch (左) 和 Walter Pitts (右)

在人工智能领域, 人工神经网络 (Artificial Neural Network, ANN) 简称神经网络, 是一种提出较早的基于仿生学理念设计和开展的人工智能研究方向。1958 年, 计算科学家 Rosenblatt 提出了由两层神经元组成的神经网络, 他将这一结构命名为感知器 (perceptron)。感知器是当时第一个有学习能力的人工神经网络。Rosenblatt

现场演示了感知器学习识别简单图像的过程，引起了轰动，这是机器学习的首次兴起。Rosenblatt 与感知器如图 1.2 所示。

1982 年，美国加州理工学院的优秀物理学家 John J. Hopfield 博士在总结已有神经网络结构和算法的基础上，提出了 Hopfield 网络模型。该模型的创新之处在于将神经网络视为一种动态系统，并运用物理力学的分析方法研究了神经网络模型的稳定性。在 1986 ~ 1988 年期间，以美国的 David E. Rumelhart 和 James L. McClelland 为代表的研究团队创作了《并行分布式处理》一书，该书基于感知器提出了多层感知器的反向传播算法（Back Propagation），简称 BP 算法。由此，神经网络从感知器结构开始，逐渐发展为包含一个隐含层的 BP 神经网络，这是机器学习的再次兴起并迎来了它的第一次浪潮（浅层学习），由此也开启了基于统计模型的机器学习方法。



图 1.2 Rosenblatt 与感知器

神经网络结构演化的历程如图 1.3 所示。

经典神经网络理论先天的缺陷是随着网络层数的增加，系统的计算量呈指数级别增长，同时前期信号经过多次衰减之后在后续层逐渐趋向为零。因此，早期的神经网络模型通常是只有一到两层隐含层的简单网络架构，无法实现带有更多隐含层的复杂网络架构。这样简单的神经网络结构无法分析较为复杂的现实问题，神经网络的研究也因此而长期停留在理论层面，鲜有实际应用的例子。

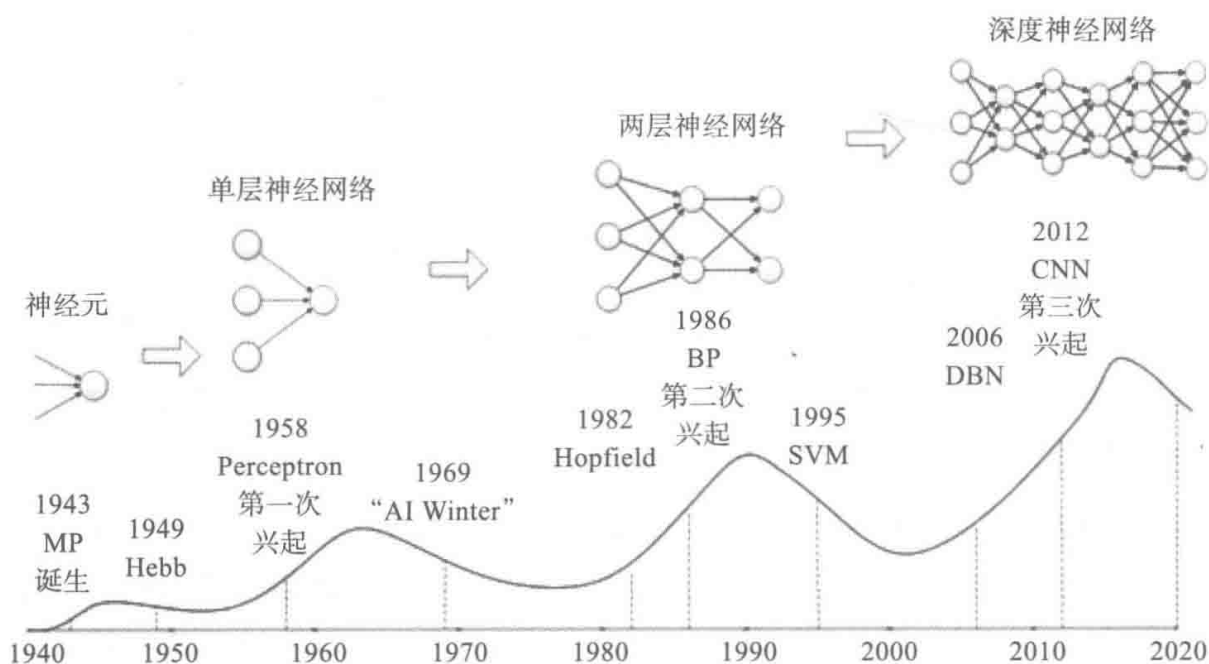


图 1.3 神经网络结构演化的历程

2006年, Geoffrey Hinton 和 Ruslan Salakhutdinov 在 *Neural Computation* 上发表论文 “A Fast Learning Algorithm for Deep Belief Nets”, 提出了深层神经网络训练中梯度消失问题的解决方案, 并给出了一种基于逐层贪心算法的深度信念网络模型。深度信念网络模型采用多个隐含层的深度结构来代替传统神经网络中感知器的单层结构, 实现了由传统的人工神经网络向深度学习的转变。由此, 机器学习迎来了它的第三次兴起和第二次浪潮, 即深度学习 (Deep-learning)。此后, 深度学习得到了飞速发展。

由于深度学习模拟人脑的分层模型结构, 具有很大的研究和应用价值, 因此, 美国国防部 (Defense Advanced Research Projects Agency, DARPA) 于 2010 年首次资助了深度学习项目, 参与项目的有斯坦福大学、纽约大学和 NEC 美国研究院。2011 年以来, 微软研究院和谷歌公司的语音识别研究人员先后采用 DNN (Deep Neural Network, 深度神经网络) 技术使语音识别错误率降低了 20% ~ 30%, 实现了语音识别领域的突破性进展, 图 1.4 所示为微软小娜语音助手。2012 年, DNN 技术在图像识别领域也取得了惊人的成果, 在 ImageNet (一个可共享的图像数据集) 评测中将错误率从 26% 降低到 15%。2016 年 3 月, 由谷歌 (公司) 旗下的 DeepMind 公司开发的 AlphaGo 与围棋世界冠军李世石进行了围棋人机大战, 并以 4 比 1 的总比分获胜, 标志着深度学习在决策相关的分支领域已经超过了人类的顶级水平。