

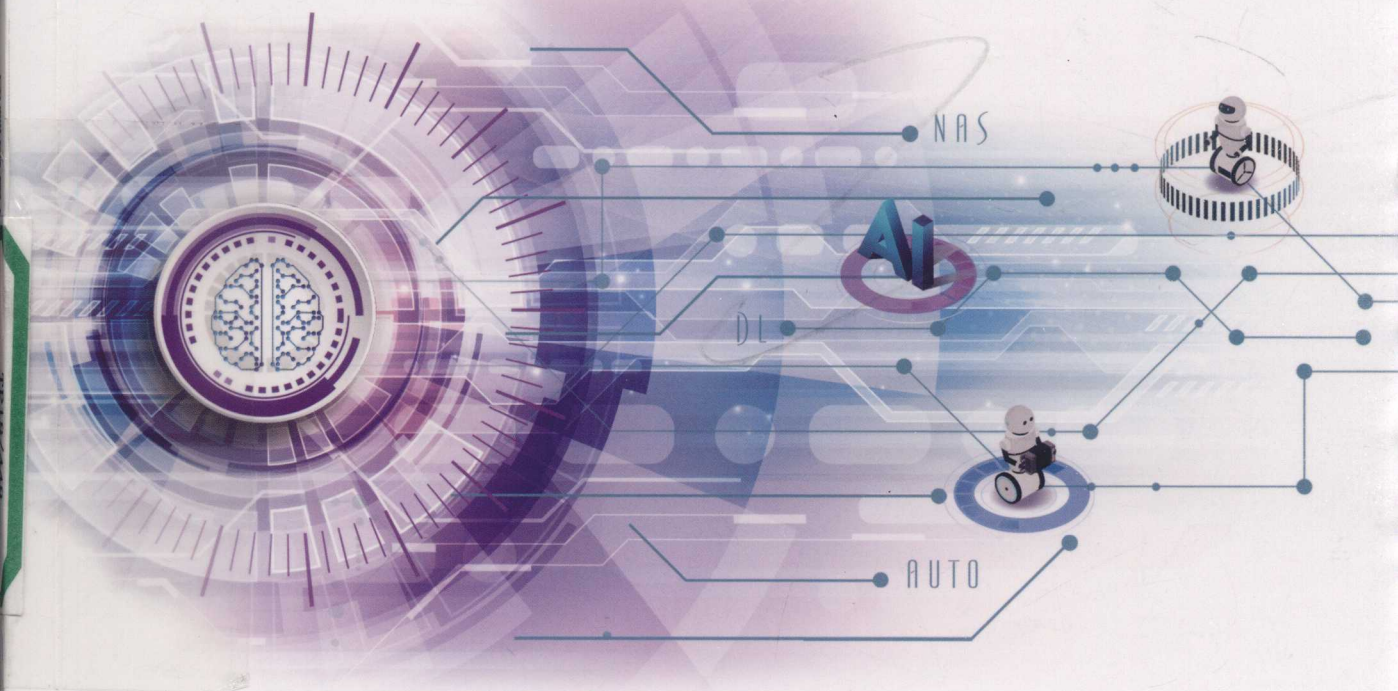
资深AI专家撰写，腾讯、阿里、微众银行、字节跳动、浙江大学、新智元等企业界、学术界、媒体界的8位资深专家联袂推荐
从基础理论、核心原理、前沿算法等多个维度全面解读AutoML、AutoDL和元学习等技术栈

Dive into AutoML and AutoDL
Building Automated Platforms for Machine Learning and Deep Learning

深入理解 AutoML和AutoDL

构建自动化机器学习与深度学习平台

王健宗 瞿晓阳 著



机械工业出版社
China Machine Press

智能系统与技术丛书

Dive into AutoML and AutoDL
Building Automated Platforms for Machine Learning and Deep Learning

深入理解 AutoML和AutoDL

构建自动化机器学习与深度学习平台

王健宗 瞿晓阳 著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

深入理解 AutoML 和 AutoDL：构建自动化机器学习与深度学习平台 / 王健宗，瞿晓阳著.
—北京：机械工业出版社，2019.8
(智能系统与技术丛书)

ISBN 978-7-111-63436-2

I. 深… II. ①王… ②瞿… III. ①人工智能—基础知识 ②机器学习—基础知识 IV. TP18

中国版本图书馆 CIP 数据核字 (2019) 第 162944 号

深入理解 AutoML 和 AutoDL

构建自动化机器学习与深度学习平台

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：罗词亮

责任校对：李秋荣

印 刷：大厂回族自治县益利印刷有限公司

版 次：2019 年 8 月第 1 版第 1 次印刷

开 本：186mm×240mm 1/16

印 张：21.75

书 号：ISBN 978-7-111-63436-2

定 价：99.00 元

客服电话：(010) 88361066 88379833 68326294

投稿热线：(010) 88379604

华章网站：www.hzbook.com

读者信箱：hzit@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

内容简介

这是一部从基础理论、核心原理、前沿算法等多个维度系统、全面讲解AutoML、AutoDL和元学习的著作。

作者是资深的人工智能专家，平安科技深度学习平台和AutoML平台负责人。本书得到了IEEE Fellow/ACM杰出科学家/香港科技大学教授杨强、腾讯AI Lab副主任俞栋、美国佛罗里达大学教授李晓林等8位来自企业界、学术界和媒体界的资深专家的一致好评。它既能让新手理清AutoML的脉络，快速上手机器学习，又能让有经验的从业者全面掌握AutoML知识体系，工作变得更高效。

全书共14章，逻辑上分为四部分：

第一部分（第1~2章）人工智能基础

对人工智能、自动化人工智能的重要概念、发展历程及现状、适用场景、主要的工具和技术等做了全面的介绍，并引出了人工智能技术未来的发展方向——AutoML，这部分是阅读本书的基础。

第二部分（第3~6章）AutoML

主要讲解机器学习和自动化机器学习，核心是AutoML，包含自动化特征工程、自动化模型选择和自动化超参优化3方面的内容。

第三部分（第7~13章）AutoDL

主要讲解深度学习和自动化深度学习，重点讲解了AutoDL的原理、基于强化学习的AutoDL、基于进化算法的AutoDL、AutoDL的高阶知识、自动化模型压缩与加速，以及各种核心算法和前沿算法。

第四部分（第14章）元学习

元学习是人工智能的理想目标，这部分对元学习的概念、流程和各种主流的学习方法都进行了详尽的介绍。

作者简介

王健宗 平安科技副总工程师、深度学习平台和AutoML平台负责人，中国人工智能开源软件发展联盟副理事长，美国佛罗里达大学人工智能博士后，曾任美国莱斯大学电子与计算机工程系研究员，专注于联邦学习和人工智能在金融、保险、投资、银行和医疗等领域的研发工作，发表联邦学习、深度学习、云计算和大数据等领域国际论文30余篇，以及发明专利200余项。多届国内知名大数据、人工智能、金融科技和联邦学习会议/论坛主席和出品人。

瞿晓阳 华中科技大学计算机系统结构博士，美国中佛罗里达大学访问学者，大型金融集团科技公司资深算法工程师，一直从事机器学习、大数据、体系结构方面的研究工作，在AutoML平台、面向AI的云原生架构、高性能计算、高效能存储系统等方面经验丰富。近几年，在国际顶级会议和顶级期刊发表过多篇文章，担任过多个国际顶级期刊的评委。



AutoML前沿

微信扫码关注，紧跟AutoML前沿

投稿通道

联系人：杨福川 邮箱：yfc@hzbook.com 微信：15693352

试读结束：需要全本请在线购买：

www.ertongbook.com

封面设计 锡彬



華章圖書

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。

PRAISE

赞 誉

自动化机器学习是未来人工智能的一个重要发展方向，值得我们关注和学习。本书作者王健宗博士是机器学习方面的资深研究者和实践者，在本书中，他不仅讲解了 AutoML 的基础理论知识，还详细分析了近几年 AutoDL 方面最前沿的算法和技术，提供了很好的方法与思路参考。想要系统研究 AutoML 并把握其最新技术趋势的读者，本书不容错过。

——杨强 IEEE Fellow/IAPR Fellow/AAAS Fellow/ACM 杰出科学家 /
微众银行首席人工智能官 / 香港科技大学计算机与工程系教授

人工智能技术将会重塑很多行业，而人工智能的自动化将极大地加速这一进程。本书作者王健宗博士是人工智能领域的知名专家，兼具深厚的理论功底和丰富的实践经验，一直致力于让 AI 无处不在。他的这本专著深入浅出地总结了人工智能自动化的基本理论、框架和技术，对研究和应用人工智能自动化的专业人士和初学者来说都是一本不可多得的参考书。

——俞栋 腾讯 AI Lab 副主任 / 西雅图人工智能实验室负责人

AutoML 可以使机器学习的调参建模流程实现自动化，大大降低机器学习的门槛，让用户在没有丰富机器学习经验的情况下也能开发机器学习模型，大大缩短创建模型的时间。王健宗博士的这本书不仅非常系统、深入地讲解了 AutoML 和 AutoDL 的理论和核心技术，而且给出了具体的工程实践方法。对于想学习 AutoML 和 AutoDL 技术或搭建自动化机器学习平台的读者来讲，本书有很大的参考价值。

——李晓林 美国佛罗里达大学教授 /
同盾科技副总裁兼人工智能研究院院长

自动化机器学习探索一种“学习的学习”模式，既是国际学术界热点研究问题，也是工业界急需的核心技术。本书是国内顶尖 AI 专家王健宗博士在实践中摸索出的自动化机器学习理论、方法与系统框架，具有极高的学术和应用价值。

——陈为 教授 / 博导 / 浙江大学计算机学院副院长

AutoML 在机器学习的发展过程中一直是研究者的梦想，让模型的超参数甚至模型结构本身就可以通过学习过程自动探索到最优解。王健宗等著的这本 AutoML 专著涵盖自动化特征工程、自动化超参优化，以及神经网络架构搜索等前沿技术方法。从理论和实践双重维度，对自动化机器学习做了全面介绍，对工程实践也有很好的指导意义。

——李磊 字节跳动人工智能实验室总监

AutoML 的概念于 2012 年由学术界提出，当时的目的是解决编程时人工调校参数的问题。在谷歌、微软等公司的大力推动下，目前 AutoML 的研究越来越深入，应用越来越广泛。AutoML 可以帮助选择模型并选择超参数，成为机器学习和 AI 自动化和平民化的重要方向。王博士的这本书填补了 AutoML 中文图书的空白，从理论与实践的双重维度，对 AutoML 和 AutoDL 的入门知识和进阶知识做了全面介绍，值得一读！

——陈继东 蚂蚁金服 ZOLOZ 全球可信身份平台负责人 / 资深数据专家

AutoML 是当前 AI 领域的前沿方向之一，由谷歌等巨头引领，让深度学习的使用更趋自动化、民主化，是降低机器学习门槛的一大利器。王健宗博士查阅并解读了近百篇论文，详尽介绍了最前沿的 AutoML 算法和技术，本书堪称打开 AutoML 和 AutoDL 深入学习之门的神奇钥匙。

——杨静 新智元创始人兼 CEO

如今，自动化机器学习技术已经风靡全球。模型选择、特征工程、调参等工作自动化之后，普通民众学习 AI 以及专业人士运用 AI 的效率得到了极大提升。王博士既是奋斗在科研一线的 AI 专家，又是平安科技的大型 AI 项目负责人，理论与实践能力均炉火纯青，写这本书再适合不过。相信很多人能从本书中受益。

——周磊 (July) 七月在线创始人兼 CEO

前 言

为什么要写这本书

“人工智能”“机器学习”“深度学习”“联邦学习”“自动化”等已经成为互联网行业使用最频繁的词汇，在人工智能发展日益成熟的今天，越来越多的研究者将目标聚焦于“自动化”。出于对 AutoML 技术出现的振奋和对人工智能的热情与投入，我们逐渐萌生了撰写这本书的想法，我们想让更多的人了解 AutoML，了解我们身边最前沿的技术和知识，最终能够让天下没有难的 AI，实现普惠 AI。如果一定要问我们写这本书的原因，我觉得可以归结为如下三点：

首先，已经有多家互联网公司发布了 AutoML 平台，毫无疑问 AutoML 已经成为目前各大公司的“护城河”，我们希望通过本书来揭开 AutoML 平台的神秘面纱。基于 AutoML 平台，专业编程人员和非专业人员均可快速创建项目并训练模型，但是，由于国内至今还没有一本关于 AutoML 算法介绍的书籍，平台用户只知其然却不知其所以然。

其次，我们想要通过本书建立一套完整的 AutoML 知识体系。很多 AutoML 从业者懂技术，但是缺少一套完整的知识体系来支撑自己的核心技术，有鉴于此，我们在开始撰写本书前做的第一件事就是建立知识体系，包括自动化机器学习、神经架构搜索的核心算法、自动化模型压缩、模型调参、深度学习的垂直领域应用以及元学习等。这套知识体系可以帮助很多从业者认清技术方向，也可以帮助初期从业者选择研究领域。我们期望有更多人来为 AutoML 这个诞生仅仅一年半的新技术添砖加瓦，共建 AutoML 生态。

最后，我们希望这本书能为更多非专业人员带来价值。本书的初期定位是 AutoML 入门书籍，换句话说，我们撰写的初衷是想为更多不懂算法但是热爱 AI 技术的爱好者提供一些思路和理解角度。因此，我们在本书中尽量使用白话来解释算法思想，从人工智能的初期发展到 AutoML 技术的成熟，可以让每一个非技术人员快速理解 AutoML。

对于本书，我们倾注了很多热情和心血，从 2017 年年底 AutoML 技术开始出现就开始

深入探索，接着起草最初书稿框架到成型历时一年多，其中经过了多次章节结构调整和修改，查阅并解读近百篇 AI 前沿论文，才有了今天大家看到了这本书。在本书中，我们从 0 到 1 介绍了 AutoML 技术的方方面面，希望这本书能带给你惊喜。

读者对象

本书适用于非计算机专业研究人员、期望转型 AI 领域的技术爱好者，同样也适用于初级、中级和高级的人工智能算法工程师、项目经理和产品经理等。

本书特色

AutoML 技术的发展日新月异，诸多科学家和研究者会在论文中发表自己的研究成果，但是目前国内还没有一本讲解 AutoML 发展和技术的书籍。本书聚焦于 AutoML，从无到有地介绍了 AutoML 的发展过程以及相关的算法。本书涉及 AutoML 技术的多个方面，从 AutoML 到 AutoDL，最后延伸到元学习，为读者提供了一套完整的知识体系。

如何阅读这本书

本书是关于自动化人工智能的一本入门级书籍，书中涵盖了大部分基础知识，因此非专业人士也可以读懂。自动化人工智能的最重要的两个分支是自动化机器学习和自动化深度学习，因此，本书的核心和聚焦在这两大研究领域，旨在为专业人士和刚入门的学者提供一些研究方向和思路。

从逻辑上，全书一共分为四个部分。

第一部分（第 1~2 章）是关于人工智能的基础概述，并介绍了现有的 AutoML 平台。

第二部分（第 3~6 章）是自动化机器学习，这里的机器学习是指统计机器学习，这一部分主要介绍了基本的机器学习知识以及自动化特征工程、自动化模型选择和自动化超参优化。

第三部分（第 7~13 章）是自动化深度学习，众所周知，近年来深度学习的研究开展得如火如荼，为了拓展读者的知识领域和研究思路，我们在这一部分花费了大量的篇幅来介绍近几年最前沿的算法和技术，这也是全书最核心的章节。

第四部分（第 14 章）是关于元学习的内容，我们认为元学习应该是独立于统计机器学习和深度学习的一个研究领域，因为元学习跳出了学习“结果”的这种思想，学习的是“学习过程”，这也是自动化人工智能的理想目标。因此，我们将元学习单独作为一个部分，作为全书内容的升华，读者可以在本书的引导下展开更深入的研究。

我们将本书的重点内容罗列为以下几点：

1) 自动化特征工程生成方法, 分别是深度特征合成算法、Featuretools 自动特征提取以及基于时序特征的自动化特征工程。

2) 自动化模型选择方法, 包括贝叶斯优化算法、进化算法、分布式优化等。

3) 自动化超参优化, 主要有序列超参优化、进化算法的运用以及迁移学习方法。

4) 神经架构搜索, 主要搜索算法有强化学习和进化算法。

5) 神经架构搜索加速方案, 包括权值共享法、超网络、网络态射法、代理评估模型以及可微分神经架构搜索。

6) 模型压缩和加速方案, 包括量化、修剪法、稀疏化以及轻量级模型设计。

专业读者或具体从业者可根据自己的研究领域以及感兴趣情况选择以上部分内容重点阅读。对于非专业读者, 本书中也有最基本的算法入门介绍, 可以将本书作为一本 AutoML 入门书籍进行全书通读。

勘误和支持

本书并没有涵盖 AutoML 研究领域的全部知识, 因为这个领域的知识体系之庞大, 不是一本书就可以介绍完的。譬如我们书中所涉及的图计算网络、超网络、蒙特卡洛树搜索以及元学习都可以成为一个独立的研究课题。在 AutoML 技术的发展过程中, 很多前沿算法会不断被提出和更新, 因此书中的内容会存在一定的局限性。

本书的很多思想和知识体系都是作者基于自己的理解建立的, 难免会出现理解不当或者不准确的地方, 恳请读者批评指正。如果你有更多的宝贵意见, 欢迎发送邮件至邮箱 yfc@hzbook.com, 我们会认真采纳你的意见和建议。这本书的结束并不意味着我们的研究就此结束, 我们还需要不断挖掘其中的精华与奥妙, 期待能够得到你们的真挚反馈和支持。

致谢

在本书的撰写和研究期间, 感谢多名 AutoML 技术爱好者 (赵淑贞、尚迪雅、曾昱为、吴文启、唐彦玺、张君婷、贺凡等) 的参与支持。

感谢出版社对本书的耐心修订和整理, 没有他们, 就没有今天这本书的出版。

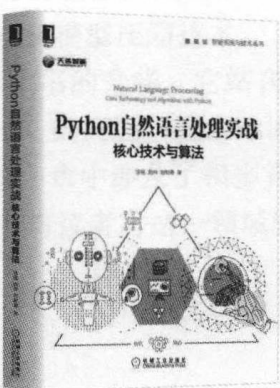
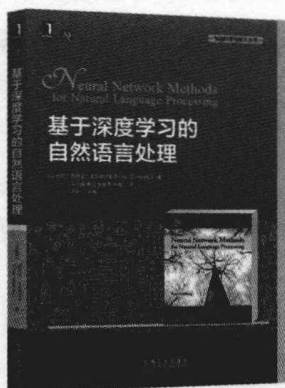
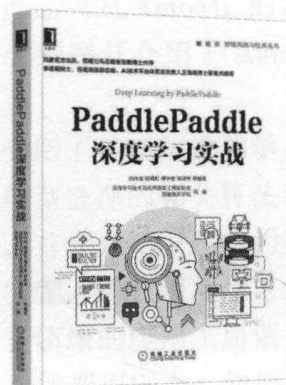
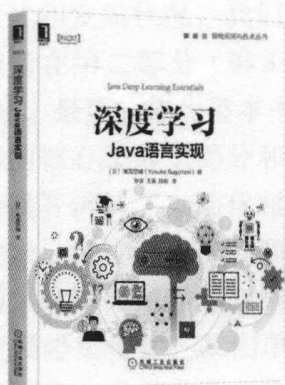
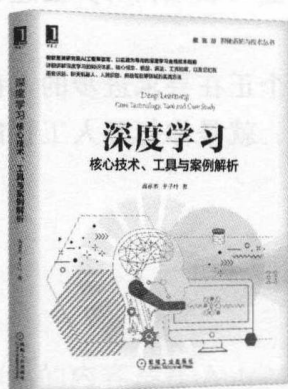
最后, 我要感谢读者, 感谢读者对我们的信任。我们尽最大努力想要给大家呈现一本逻辑清晰、技术易懂的入门书籍, 感谢读者选择了这本书, 选择就是对我们最大的信任。

谨以此书献给 AutoML 的技术爱好者和研究者们!

王健宗

2019 年 8 月于深圳

推荐阅读



CONTENTS

目 录

赞誉		
前言		
第 1 章 人工智能概述	1	
1.1 全面了解人工智能	1	
1.1.1 人工智能定义	1	
1.1.2 弱人工智能、强人工智能与 超人工智能	2	
1.1.3 人工智能三大主义	3	
1.1.4 机器学习与深度学习	4	
1.2 人工智能发展历程	5	
1.3 深度学习的崛起之路	7	
1.3.1 人脸识别的起源	7	
1.3.2 自动驾驶的福音	7	
1.3.3 超越人类的 AI 智能体	8	
1.3.4 懂你的 AI	8	
1.3.5 奔跑、飞行以及玩游 戏的 AI	8	
1.3.6 人人都可以创造属于自 己的 AI	8	
1.4 深度学习的发展	9	
1.4.1 计算机视觉	9	
1.4.2 自然语言处理	10	
1.4.3 语音识别	11	
1.5 下一代人工智能	11	
1.6 参考文献	13	
第 2 章 自动化人工智能	14	
2.1 AutoML 概述	14	
2.1.1 什么是自动化	14	
2.1.2 AutoML 的起源与发展	15	
2.2 AutoML 的研究意义	17	
2.2.1 AutoML 的研究动机	17	
2.2.2 AutoML 的意义和作用	18	
2.3 现有 AutoML 平台产品	21	
2.3.1 谷歌 Cloud AutoML	21	
2.3.2 百度 EasyDL	23	
2.3.3 阿里云 PAI	24	
2.3.4 探智立方 DarwinML	28	
2.3.5 第四范式 AI Prophet AutoML	29	
2.3.6 智易科技	30	
2.4 参考文献	31	

第 3 章 机器学习概述	32	4.6 自动化特征工程工具	67
3.1 机器学习的发展	32	4.6.1 自动化特征工程系统	67
3.1.1 “机器学习”名字的由来	32	4.6.2 自动化特征工程平台	71
3.1.2 “机器学习”的前世今生	33	4.7 参考文献	75
3.1.3 “机器学习”的理论基础	34	第 5 章 自动化模型选择	76
3.2 机器学习的实现方法	36	5.1 模型选择	76
3.2.1 分类问题	36	5.2 自动化模型选择	77
3.2.2 回归问题	38	5.2.1 基于贝叶斯优化的自动化 模型选择	78
3.2.3 聚类问题	39	5.2.2 基于进化算法的自动化模 型选择	84
3.3 自动化机器学习	40	5.2.3 分布式自动化模型选择	86
3.3.1 机器学习面临的问题	40	5.2.4 自动化模型选择的相关 平台	92
3.3.2 为什么会产生 AutoML	41	5.3 自动集成学习	94
3.4 参考文献	41	5.3.1 集成学习基础	94
第 4 章 自动化特征工程	43	5.3.2 集成学习之结合策略	97
4.1 特征工程	43	5.3.3 自动化模型集成	98
4.1.1 什么是特征	43	5.4 参考文献	99
4.1.2 什么是特征工程	44	第 6 章 自动化超参优化	101
4.2 特征工程处理方法	45	6.1 概述	101
4.2.1 特征选择	45	6.1.1 问题定义	103
4.2.2 数据预处理	47	6.1.2 搜索空间	103
4.2.3 特征压缩	48	6.1.3 搜索策略	103
4.3 手工特征工程存在的问题	49	6.1.4 评价预估	104
4.4 自动化特征工程	50	6.1.5 经验迁移加速	105
4.4.1 什么是自动化特征工程	50	6.2 基本方法	105
4.4.2 机器学习和深度学习的 特征工程	51	6.2.1 网格搜索	105
4.5 自动化特征工程生成方法	52	6.2.2 随机搜索	105
4.5.1 深度特征合成算法	52	6.3 基于模型的序列超参优化	106
4.5.2 Featuretools 自动特征提取	52		
4.5.3 基于时序数据的自动化 特征工程	56		

6.3.1	代理模型的选择	108	7.3	CNN 经典模型	126
6.3.2	代理模型的更新	108	7.3.1	LeNet	126
6.3.3	新超参组的选择	109	7.3.2	AlexNet	127
6.3.4	基于高斯过程回归的序列超参优化	111	7.3.3	VGGNet	128
6.3.5	基于随机森林算法代理的序列超参优化	112	7.3.4	GoogLeNet	129
6.3.6	基于 TPE 算法的序列超参优化	114	7.3.5	ResNet	130
6.3.7	SMBO 的进阶技巧	114	7.3.6	DenseNet	131
6.4	基于进化算法的自动化超参优化	115	7.4	循环神经网络	132
6.4.1	基于进化策略的自动化超参优化	115	7.4.1	基本循环神经模型	132
6.4.2	基于粒子群算法的自动化超参优化	116	7.4.2	LSTM 模型	133
6.5	基于迁移学习的超参优化加速方法	117	7.4.3	GRU 模型	134
6.5.1	经验迁移机制	117	7.5	参考文献	134
6.5.2	经验迁移衰退机制	117			
6.5.3	经验迁移权重机制	117	第 8 章	自动化深度学习概述	136
6.5.4	优化过程的试点机制	118	8.1	深度学习 vs 自动化深度学习	136
6.6	参考文献	118	8.2	什么是 NAS	136
			8.2.1	问题定义	137
第 7 章	深度学习基础	120	8.2.2	搜索策略	139
7.1	深度学习简介	120	8.2.3	加速方案	140
7.1.1	什么是神经元	120	8.3	NAS 方法分类	140
7.1.2	人工神经网络的发展历程	121			
7.1.3	深度学习方法	123	第 9 章	基于强化学习的 AutoDL	142
7.2	卷积神经网络简介	123	9.1	强化学习基础	142
7.2.1	卷积层	123	9.1.1	强化学习简介	142
7.2.2	池化层	125	9.1.2	基本要素及问题定义	144
7.2.3	全连接层	126	9.1.3	发展历史	144
			9.1.4	基本方法	146
			9.2	两类基本模型	147
			9.2.1	TD 经典算法	148
			9.2.2	DQN 系列算法	149
			9.2.3	策略梯度算法	152
			9.3	强化学习之 Actor-Critic 系列	154

9.3.1 Actor-Critic 算法	154	10.3.2 群体算法——以 PSO 为例	205
9.3.2 确定性策略梯度	155	10.3.3 文化基因算法	207
9.3.3 深度确定性策略梯度	157	10.3.4 差分进化算法	208
9.3.4 异步优势 Actor-Critic 算法	158	10.3.5 分布估计算法	208
9.3.5 近端策略优化	160	10.4 进化神经网络	209
9.3.6 分布式近端策略优化	164	10.4.1 简介	209
9.4 基于强化学习的自动搜索	166	10.4.2 神经网络编码方式	210
9.5 基本搜索方法	166	10.4.3 竞争约定	211
9.5.1 基于层的搜索	166	10.4.4 网络结构的创新性	212
9.5.2 基于块的搜索	169	10.4.5 NAS 之进化算法	212
9.5.3 基于连接的搜索	171	10.5 细粒度的神经进化 (NEAT 算法)	213
9.6 进阶搜索方法	173	10.5.1 基因编码	214
9.6.1 逆强化学习	173	10.5.2 基因的可追溯性	216
9.6.2 图超网络	174	10.5.3 通过物种形成保护创新结构	216
9.6.3 蒙特卡洛树搜索	175	10.6 粗粒度的神经进化 (CoDeep-NEAT 算法)	218
9.6.4 知识提炼 (教师网络)	177	10.6.1 DeepNEAT 算法	218
9.7 参考文献	179	10.6.2 CoDeepNEAT 算法	219
第 10 章 基于进化算法的 AutoDL	181	10.7 block-level 的进化	220
10.1 启发式算法	181	10.7.1 Genetic CNN 算法	220
10.1.1 随机搜索	182	10.7.2 CGP-CNN 方法	222
10.1.2 近邻搜索	183	10.8 基于 node-level 的网络架构进化	224
10.1.3 进化计算	187	10.8.1 思想简介	224
10.1.4 启发式算法的局限性	189	10.8.2 基本算法设计	225
10.2 初代进化算法	190	10.8.3 信息复用与加速	226
10.2.1 基本术语	190	10.9 基于 NAS 搜索空间的网络架构进化	227
10.2.2 基础算子	191		
10.2.3 遗传算法	196		
10.2.4 进化策略	198		
10.2.5 进化规划	199		
10.3 其他近代进化算法	200		
10.3.1 遗传编程算法簇	200		

10.9.1	思想简介	227	第 12 章 垂直领域的 AutoDL	252	
10.9.2	基本算法设计	227	12.1	AutoCV	252
10.9.3	信息复用与加速	228	12.1.1	Auto-DeepLab (图像语义分割)	252
10.10	基于层次拓扑表示的网络进化方法	228	12.1.2	随机连线神经网络	257
10.10.1	思想简介	228	12.2	AutoVoice	261
10.10.2	分级表示	229	12.2.1	关键词定位问题定义	261
10.10.3	随机的层次分级进化	230	12.2.2	随机自适应架构搜索原理	262
10.11	参考文献	230	12.2.3	SANAS 模型	262
第 11 章 AutoDL 高阶		233	12.3	AutoNLP	263
11.1	搜索加速之权值共享法	233	12.3.1	什么是自注意力机制	263
11.1.1	ENAS	233	12.3.2	初识 Transformer 模型	265
11.1.2	基于稀疏优化的 NAS	235	12.3.3	Evolved Transformer 结构	266
11.2	基于 one-shot 模型的架构搜索	236	12.4	参考文献	270
11.2.1	超网络的应用	236	第 13 章 自动化模型压缩与加速	271	
11.2.2	基于 one-shot 的搜索	237	13.1	从生物角度看模型压缩的重要性	271
11.2.3	实例级架构搜索	238	13.1.1	人脑神经元的修剪	271
11.2.4	单路径超网络	240	13.1.2	大脑的冗余性	272
11.3	搜索加速之代理评估模型	241	13.1.3	修剪的意义	273
11.3.1	代理模型	241	13.2	模型压缩发展概述	274
11.3.2	PNAS 中的 LSTM 代理	242	13.3	入门级方法: 量化技术	275
11.4	基于网络态射法的神经架构搜索	244	13.3.1	量化技术	275
11.4.1	网络态射的提出	244	13.3.2	二值化网络	276
11.4.2	什么是网络态射	244	13.3.3	TensorRT	277
11.4.3	网络态射 + 迂回爬山法	246	13.4	初级方法: 修剪法	278
11.5	可微分神经架构搜索	247	13.4.1	修剪法	278
11.5.1	可微分神经架构搜索的来源	247	13.4.2	修剪与修复	279
11.5.2	可微分神经架构搜索的方法	248	13.5	中级方法: 稀疏化技术	281
11.6	参考文献	250	13.5.1	正则化	281
			13.5.2	知识精炼	281