





How Smart Machines Think



智能机器 如何思考

[美] 肖恩·格里什 (Sean Gerrish) 著
张羿 译

中信出版集团

 智能机器 
 如何思考 

[美] 肖恩·格里什 (Sean Gerrish) 著 张羿 译

**How Smart
Machines Think**

图书在版编目 (CIP) 数据

智能机器如何思考 / (美) 肖恩·格里什著; 张羿
译. -- 北京: 中信出版社, 2019.7
书名原文: How Smart Machines Think
ISBN 978-7-5217-0546-1

I. ①智… II. ①肖… ②张… III. ①信息经济-研
究②人工智能-研究 IV. ①F49 ②TP18

中国版本图书馆CIP数据核字 (2019) 第 086106 号

How Smart Machines Think by Sean Gerrish
Copyright © 2018 Massachusetts Institute of Technology
Published by arrangement with The MIT Press through Bardon-Chinese Media Agency
Simplified Chinese translation copyright © 2019 by CITIC Press Corporation
ALL RIGHTS RESERVED

本书仅限中国大陆地区发行销售

智能机器如何思考

著 者: [美] 肖恩·格里什
译 者: 张 羿

出版发行: 中信出版集团股份有限公司
(北京市朝阳区惠新东街甲4号富盛大厦2座 邮编 100029)

承 印 者: 北京画中画印刷有限公司

开 本: 880mm × 1230mm 1/32 印 张: 13.25 字 数: 282千字
版 次: 2019年7月第1版 印 次: 2019年7月第1次印刷
京权图字: 01-2018-7852 广告经营许可证: 京朝工商广字第8087号
书 号: ISBN 978-7-5217-0546-1
定 价: 69.00元

版权所有·侵权必究
如有印刷、装订问题, 本公司负责调换。
服务热线: 400-600-8099
投稿邮箱: author@citicpub.com

献给设计和构建智能机器的工程师和研究员

推荐序

10多年前，我遇见了肖恩。当时我在谷歌领导团队，负责开发为谷歌的搜索广告业务提供支持的许多大型机器学习系统。肖恩是我们小组里最顶尖的工程师之一，当时他正在研究机器学习前沿领域的一系列具有挑战性的问题。我们一起工作以来，体现在统计机器学习技术中的各类人工智能已经从相对难以触及的神秘技术、研究人员和高科技公司的专属领域，发展成为日益平易近人的、卓有成效的工具和技术，值得每一位软件开发人员使用。

目前机器学习领域取得的快速进展，在一定程度上是由以下因素推动的：数据爆炸，高性能计算机体系结构的复兴，云提供商竞相为开发人员和研究人员构建可扩展的人工智能平台，人们将实时智能嵌入移动设备、汽车、其他消费电子产品和日益普遍的连接到云端的计算设备的热潮。这种快速进步中包括一些惊世骇俗的成就，机器在许多狭窄的领域已经接近于或超过了人类的能力，例如在图像中标记物体、识别语音、玩策略游戏以及翻译

语言，但我们仍处于这些技术发展的初期，摆在我们面前的是长达几十年的创新和发现之旅。

对开发人员和研究人员而言，理解机器学习的工作原理是一个明智的职业选择。目前，全球的科技巨头公司对这些技术的专业知识都有很高的需求。微软、亚马逊、谷歌、苹果、百度等许多公司都提供应用程序接口、工具包和云计算基础设施，将机器学习的开发工作交给全世界数以千万计的开发人员。未来几年，随着越来越多的应用程序包含智能功能，大多数开发人员都需要掌握一些机器学习技术。这正是这本书的宝贵价值所在。

这本书诞生自肖恩对了解现代机器学习成功之路的渴望。在用清晰易懂的方式描述这些系统的本质时，肖恩利用 10 多年的行业和学术经验解决了机器学习带来的一些最棘手的问题。鉴于机器学习系统能够复制某些方面的人类智能，等到某个突破点临近，原本由人类特有的创造诗意词句的能力或许也会被机器复制。肖恩对这些技术严谨实用的描述反映了他在科研战壕中的岁月，不时令人感到痛苦的反复试验让战壕中的人们了解到，机器学习并不是魔法。如果你知道如何应用它，了解它的局限所在，它就是强有力的前沿工具；如果你不知道，它就几乎一文不值。

肖恩通过列举现实世界的例子，回避不必要的术语，使现代机器学习的概念变得通俗易懂。这本书假设读者在机器学习或计算机科学领域的知识相对较少，因此对更广泛的受众而言非常友好。鉴于当前围绕机器学习和人工智能的对话颇为活跃，并且这些技术可能对我们的未来产生影响，任何想要参与这场对话的人

都应该尽可能地学习。由于市场上缺乏对机器学习通俗易懂的专业介绍，这本书将成为引导你理解底层技术的理想方法，它可以帮助你更好地判断哪些言论值得相信，哪些说得天花乱坠的谬论应该被抛弃。

微软CTO（首席技术官）

凯文·斯科特（Kevin Scott）

前 言

2010 年的一个晚上，这本书的种子播种在了计算机科学系顶层的人工智能研究实验室里。当时，我刚刚参加了一些关于自动驾驶汽车的研讨会，对它们的工作原理颇为好奇，于是我做了一些网络搜索。我找到的最好的解释是卡内基-梅隆大学和斯坦福大学的一些研究人员撰写的学术论文。我看了几分钟，对自动驾驶汽车的工作原理有了些粗浅的认识，然后便离开了。

但随着时间的推移，我发现自己三番五次地重复这个过程。每当看到媒体报道人工智能或机器学习领域的一个个突破时，我就会回到同一个问题：它们是如何工作的？令我感到奇怪的是，我花了无数时间在学术界和产业界研究和实践机器学习，但我仍然不能坚定连贯地回答这个问题。我想，或许我对人工智能和机器学习的了解不如我本应了解的那么多，又或许大学课程没有教授我们全面的知识。大多数关于这些主题的大学课程只教授这些突破背后的构件，而不教授如何将这些构件组合在一起去做有趣的事情。

但还有另一个更根本的原因，即我无法弄清楚它们的工作原理：这些突破中的大多数确实涉及开创性的研究，我们根本不知道如何构建它们，直到研究人员找到了方法、撰写了过程或构建了原型。这就是为什么研究人员一直在同行评议期刊上发表关于这些突破的文章，因为这些文章新颖、有影响力、非显而易见（而且是经过同行评议的）。但是，这些突破背后的细节一经发表就会被随意散布在许多不同的来源中，因此仍然无济于事。

最终，我意识到应该把自己在研究中学到的东西与他人分享，这样他们就无须为了理解相同的东西而跨越同样的障碍。换言之，我写这本书的原因是，这是一本如果我不懂机器学习我将会想读的书。

于是我写了这本书，希望它能够帮助那些对广义上的科学技术感兴趣的人，无论老少；或者那些想要更多地了解机器学习和人工智能是否会对他们的公司有帮助的行业领导者。这本书旨在让广大读者都能读懂，无论是满怀好奇心的高中生，还是退休的机械工程师。虽然了解一些计算机科学会有所帮助，但阅读本书唯一的前提条件是好奇心和一点专注力。我有意把这本书中涉及的数学知识保持在最低限度，以便向普通读者更好地传达核心思想。

机器人、人工智能和机器学习领域的专家通常比较了解我将要描述的一些算法的实现细节，但是对他们中的许多人而言，其余的叙述和对整个系统的设计可能仍然是陌生的（除非这是他们的研究领域）。我希望这本书能为大家带来一些新东西。

目 录

推荐序 // IX

前 言 // XIII

1 / 自动机的秘密

长笛演奏者 // 003

今天的自动机 // 005

钟摆的摆动 // 007

并不难懂的自动机 // 008

2 / 自动驾驶汽车：挑战不可能

沙漠中的百万美元竞赛 // 015

如何打造自动驾驶汽车 // 017

规划路径 // 021

路径搜索 // 022

导航 // 025

无人车挑战赛的获胜者 // 028

一场失败的比赛 // 031

3 / 保持在车道内行驶：自动驾驶汽车的感知

第二次无人车挑战赛 // 035

自动驾驶汽车中的机器学习 // 037

斯坦利的架构 // 038

避开障碍物 // 040

寻找道路的边缘 // 043

开眼看路 // 045

路径规划 // 047

斯坦利大脑的各个部分如何相互交流 // 049

4 / 在十字路口避让：自动驾驶汽车的大脑

城市挑战赛 // 055

感知抽象 // 057

比赛 // 059

Boss的高层次推理层 // 061

攻克交通堵塞 // 068

三层架构 // 071

自动驾驶汽车看到的物体 // 076

自动驾驶汽车：复杂的系统 // 077

自动驾驶汽车的轨迹 // 078

5 / 网飞和推荐引擎的挑战

百万美元大奖 // 083

竞争者 // 084

如何训练分类器 // 085

比赛的目标 // 089

庞大的评分矩阵 // 091

矩阵分解 // 096

第一年结束 // 102

6 / 团队融合：网飞奖的赢家

缩小竞争者之间的差距 // 107

第一年末 // 108

随时间变化的预测 // 111

过度拟合 // 114

模型混合 // 115

第二年 // 119

最后一年 // 120

赛后 // 124

7 / 用奖励教导计算机

DeepMind玩雅达利游戏 // 129

强化学习 // 132

教导智能体 // 134

为智能体编写程序 // 136

智能体如何观察环境 // 140

经验金块 // 141

用强化学习玩雅达利游戏 // 148

8 / 如何用神经网络攻克雅达利游戏

- 神经信息处理系统 // 153
- 近似，而非完美 // 153
- 用作数学函数的神经网络 // 155
- 雅达利游戏神经网络的结构 // 161
- 深入研究神经网络 // 170

9 / 人工神经网络的世界观

- 人工智能的奥秘 // 175
- 国际象棋自动机“土耳其人” // 177
- 神经网络中的误差 // 179
- 识别图像中的物体 // 180
- 过度拟合 // 183
- ImageNet // 185
- 卷积神经网络 // 188
- 为什么是深度神经网络？ // 194
- 数据瓶颈 // 197

10 / 深入了解深度神经网络的内部秘密

- 计算机生成图片 // 203
- 压缩函数 // 204
- ReLU激活函数 // 207
- 机器人之梦 // 211

11

能听、能说、能记忆的神经网络

对机器而言，“理解”意味着什么？ // 221

深度语音识别系统 // 222

循环神经网络 // 223

为图像生成字幕 // 230

长短时记忆网络 // 233

对抗数据 // 235

12

理解自然语言

是宣传噱头，还是人工智能研究的福音？ // 239

IBM的“沃森” // 240

攻克《危险边缘》所遇到的挑战 // 241

浩如烟海的知识 // 242

《危险边缘》挑战赛的诞生 // 245

DeepQA // 247

问题分析 // 249

“沃森”如何解读句子？ // 252

13

挖掘《危险边缘》的最佳答案

地下室基准 // 261

生成候选答案 // 263

查找答案 // 266

轻量级过滤器 // 269

证据检索 // 270

评分 // 274

汇总和排名 // 277

调整“沃森” // 281

重新审视DeepQA // 282

“沃森”有智能吗? // 283

14

用蛮力搜索找到好策略

通过搜索玩游戏 // 289

数独 // 290

树的大小 // 294

分支因子 // 297

游戏中的不确定性 // 297

克劳德·香农与信息论 // 302

评价函数 // 303

“深蓝” // 308

加入IBM // 310

搜索与神经网络 // 311

西洋双陆棋程序 // 313

搜索的局限 // 315

15

职业水平的围棋

计算机围棋 // 319

围棋 // 321

通过抽样走子来建立直觉 // 324

神之一手 // 330

蒙特卡洛树搜索 // 333

单臂老虎机 // 337

AlphaGo是否需要如此复杂 // 339

AlphaGo的局限 // 341

16 / 实时人工智能与《星际争霸》

构建更好的游戏机器人 // 345

《星际争霸》与人工智能 // 346

简化游戏 // 348

实用《星际争霸》机器人 // 351

OpenAI与《DOTA 2》 // 354

《星际争霸》机器人的未来 // 357

17 / 50年后或更遥远的未来

人工智能起起伏伏的发展过程 // 363

如何复制这本书中的成功 // 364

数据的普遍使用 // 368

下一步去向何方 // 369

致 谢 // 373

注 释 // 375



—
自动机的秘密

