



# 深度学习商业应用开发指南

——从对话机器人到医疗图像处理

Armando Vieira

Bernardete Ribeiro

张向东 江超

编著

译



北京航空航天大学出版社  
BEIHANG UNIVERSITY PRESS

Apress®

# 深度学习商业应用开发指南

——从对话机器人到医疗图像处理

Armando Vieira      编著  
Bernardete Ribeiro  
张向东 江超 译

北京航空航天大学出版社

## 图书在版编目(CIP)数据

深度学习商业应用开发指南：从对话机器人到医疗  
图像处理 / (瑞典)阿曼多·维埃拉, (葡)伯纳德·里  
贝罗编著; 张向东, 江超译. -- 北京: 北京航空航天  
大学出版社, 2019. 7

书名原文: Introduction to Deep Learning  
Business Applications for Developers: From  
Conversational Bots in Customer Service to Medical  
Image Processing

ISBN 978 - 7 - 5124 - 3039 - 6

I. ①深… II. ①阿… ②伯… ③张… ④江… III.  
①机器学习—研究 IV. ①TP181

中国版本图书馆 CIP 数据核字(2019)第 138935 号

版权所有, 侵权必究。

## 深度学习商业应用开发指南——从对话机器人到医疗图像处理

Armando Vieira 编著  
Bernardete Ribeiro  
张向东 江超 译  
责任编辑 杨昕

\*

北京航空航天大学出版社出版发行

北京市海淀区学院路 37 号(邮编 100191) <http://www.buaapress.com.cn>

发行部电话:(010)82317024 传真:(010)82328026

读者信箱: [copyrights@buaacm.com.cn](mailto:copyrights@buaacm.com.cn) 邮购电话:(010)82316936

涿州市新华印刷有限公司印装 各地书店经销

\*

开本: 710×1 000 1/16 印张: 15 字数: 311 千字

2019 年 8 月第 1 版 2019 年 8 月第 1 次印刷

ISBN 978 - 7 - 5124 - 3039 - 6 定价: 59.00 元

若本书有倒页、脱页、缺页等印装质量问题, 请与本社发行部联系调换。联系电话: (010)82317024

First published in English under the title

**Introduction to Deep Learning Business Applications for Developers: From Conversational Bots in Customer Service to Medical Image Processing**

by **Armando Vieira and Bernardete Ribeiro**

Copyright © Armando Vieira, Bernardete Ribeiro, 2018

This edition has been translated and published under licence from APress Media, LLC, part of Springer Nature.

本书中文简体字版由 APress Media, LLC, part of Springer Nature 授权北京航空航天大学出版社在中华人民共和国境内独家出版发行。版权所有。

北京市版权局著作权合同登记号 图字:01-2018-7320 号

## 译者序

1992年离开清华园，也离开了当时很有产业应用前景的通信行业，进入了语音识别这个前沿研究领域，没考虑太多的就业前景，因为那时没听说过互联网，更别说人工智能产业了，只是觉得语音识别更有趣，在智力上的挑战也更大。

27年后，刚刚在一家国内上市公司完成一个“人工智能、机器换人”项目后，就辞去这家上市公司的创新研究院院长职位，加入一家“人工智能+音乐教育”的创业公司。回顾这些年的所见所闻，感慨万千。

刚开始，语音识别和图像处理是两个不同的行当，模式识别和专家系统也是相距比较远的门类，那时很少有人把它们统称为人工智能，因为实在太难了，每个细分行当的工具差别都很大，很少有人能同时精通这几个门类。经历了近60年艰苦跋涉后，很多“AI人才”转行进入互联网、基因测序和电子产品研发领域。2012年后，深度神经网络技术借助GPU和互联网大数据，在语音识别和图像识别等领域首度超越了人的识别能力，人工智能产业的革命真正到来了！

常常有人问我：“有了人工智能会怎样？”我建议他这样思考：“在1900年，人们会问有了电会怎样？站在21世纪的你该怎样回答？”20世纪这一百年最重要的技术要素是电，人们现在不会再大规模从事“洗衣工”这样的职业。因为有了电，无论是黄昏还是黑夜，人们都可以工作、学习、娱乐，产生了不计其数的新职业，生活品质大大提升。

21世纪的最大技术要素无疑是人工智能，生产线工人、司机或物流人员、保安等职业将逐渐消失，就像100年前的洗衣工人，同时又有难以置信的无数新行业被“发明”出来，21世纪末生活的精彩程度是现在难以想象的。

无论你现在有什么技能，从事什么行业，在21世纪的生存发展都要求你具备“人工智能场景思维”——在各个场景里，人工智能做什么，人做什么，如何分工配合。这就是我们选择翻译推出这本《深度学习商业应用开发指南——从对话机器人到医疗图像处理》的原因。本书涉及的场景包括图像分割、图像识别、图像标题、视觉问答、视频分析、卫星图像处理、知识图谱、自然语言翻译、多模态学习、语音识别、机器

人控制、自动驾驶、对话机器人、电子商务推荐算法、棋类游戏、电子游戏、图画风格转化、音乐处理、信用卡防诈骗、金融预测、医疗图像识别、新药发现、法务、数据中心管理等几十种应用。围绕这些应用场景,对具体的算法和技术也做了详细讨论。

本书适合各级各类企业管理者、产品经理、软硬件工程师、测试人员阅读,也适合人工智能培训班、大学生创新创业实战训练、研究生课题演练、程序员实力提升使用。

本书的出版得益于北京航空航天大学出版社的推荐以及北航软件学院研究生邱国庆、李文意的辛勤付出,在此一并表示感谢。

由于人工智能是近年来快速发展、迭代演变的领域,对于一些术语也有不同的译法,我们尽量多方考证,选用在国内最为普遍使用的术语。

两位译者对全书做了三遍审读,尽管如此,错误仍会在所难免,如果读者发现错误或不妥之处,可以给我们发邮件,我们将在再版时修订,感谢您的参与和指正。

衷心祝愿您能拥抱人工智能时代,具备人工智能场景思维,更好地服务他人,为社会创造更多财富,也让您和您的家人拥有更美好的未来。

译者

2018年6月于苏州科技城

译者邮箱:1589492540@qq.com 或 speech99@gmail.com。

致我的家人。

——伯纳德·里贝罗 (Bernardete Ribeiro)

## 关于作者

阿曼多·维埃拉(Armando Vieira)于1997年在葡萄牙科英布拉大学(University of Coimbra)获得物理学博士学位,然后开始从事人工神经网络研究工作。他于2003年率先开展了深度神经网络研究,目前在从事图像处理、药物研发、信用评分及风险分析的公司和初创团队中担任高级数据科学顾问。他曾参与了许多与人工智能相关的商业活动,并且是 Alea.ai 的创始人。读者可以在 <http://armando.lidinwise.com> 上找到更多关于他的信息。

伯纳德·里贝罗(Bernardete Ribeiro)是葡萄牙科英布拉大学(University of Coimbra)的教授,主要教授编程、模式识别、商业智能和其他相关课程。她获得了科英布拉大学信息学工程(CISUC)博士学位,并且是 CISUC 的特许教授。同时,她还担任 CISUC 信息与系统中心的主任。她的主要研究方向包括机器学习、模式识别、金融工程、文本分类和信号处理,及其在相关领域的广泛应用。她是科英布拉大学人工神经网络实验室(LARN)的创始人,并且担任该实验室主任超过20年。伯纳德不但是葡萄牙模式识别协会(APRP)的主席,而且还是国际模式识别协会(IAPR)理事会的成员。

# 关于技术评审人

乔乔·姆拉伊(Jojo Moolayil)是一位具有五年以上人工智能、深度学习、机器学习和决策科学工业应用经验的专业人士。他是 *Smarter Decisions: The Intersection of IoT and Decision Science* (《物联网和决策科学的交叉》) 的作者,与多个数据科学、机器学习领域具有较高影响力的行业领导者开展了跨行业的合作。目前其就职于通用电气公司(GE),生活在有印度硅谷之称的班加罗尔。

乔乔出生于印度浦那(Pune),并在那里长大,毕业于浦那大学(University of Pune),主修信息技术工程。他的职业生涯始于世界上最大的数据分析公司 Mu Sigma Inc.,为世界 50 强客户提供专业的数据分析服务。乔乔·姆拉伊作为早期物联网分析的参与者,不仅将其在决策科学的知识和经验应用于问题的解决和架构的搭建上,还从数据决策科学和物联网分析的应用中获取了更多的知识和经验。

为了巩固他的数据科学研究在工业物联网应用方面的基础,扩大他在解决实验方面的影响(实验以解决问题为目的),他加入了名为 Flutura 的物联网分析初创公司,该公司的基地位于班加罗尔,总部位于加州硅谷。在与 Flutura 的短暂合作后,乔乔继续与通用电气公司在工业物联网领域进行合作,专注于解决工业物联网的决策问题。乔乔在通用电气公司的另一部分工作是为工业物联网开发数据科学和决策科学的产品和平台。

乔乔除了撰写有关决策科学和物联网方面的书籍外,还是 Apress 和 Packt 出版社出版的各种学习、深度学习和商业分析等方面书籍的技术评审员。他是一位活跃的数据科学导师,并在 [www.jojomoolayil.com/web/blog/](http://www.jojomoolayil.com/web/blog/) 上开设了博客。读者可以通过 <https://www.linkedin.com/in/jojo62000> 与他联系。

# 致 谢

感谢所有为本书做出贡献的人们,感谢他们的帮助、支持和建议。特别感谢来自 ContextVision AB 公司的马丁·赫德伦德和迈克尔·罗森(Mikael Rousson)的支持和鼓励,正是基于他们鼓舞人心的谈话,才有了此书的写作和出版。

还要感谢科英布拉大学信息与系统中心(CISUC)以及科英布拉大学信息工程系和科学技术学院,他们为本书的写作提供了大量的资源和有力的支持。

还要感谢尼尔·洛佩兹(Noel Lopes),他审阅了本书多核处理的相关部分;并感谢本杰明·阿弗斯(Benjamin Auffarth)仔细阅读了手稿并提出了相关修改建议。

特别感谢 Springer 出版社的两位编辑塞莱斯廷·约翰(Celestin John)和迪夫亚·莫迪(Divya Modi),感谢他们的鼓励。

最后,感谢我们的家人和朋友的爱和支持。

阿曼多·维埃拉(Armando Vieira)

伯纳德·里贝罗(Bernardete Ribeiro)

2018年2月于科英布拉,葡萄牙

# 序 言

深度学习已经风靡人工智能领域,几乎渗透到各个商业应用当中。由于现在几乎所有内容和交易都以数字格式记录,因此可通过机器学习算法研究大量数据。然而,通过传统的机器学习技术很难探索这种所谓的大数据中出现的错综复杂的关系,对于诸如图像、语音和文本之类的非结构化数据,尤其困难。

深度学习算法具有非常强的学习能力,可以应对分析巨大数据流的挑战。此外,深度神经网络相对于其他人工智能技术,需要很少(如果有的话)的特征工程,就可以从头到尾进行训练。深度学习算法的另一个优点是仅需要很少的监督架构(换句话说,这些架构可以自动从数据中学习,几乎不需要人为干预)。这些架构是弱监督学习,即所谓“无监督”。最后,深度学习可以当作生成过程进行训练,其算法不是将输入映射到输出,而是学习如何从纯噪声(即生成对抗网络)生成输入和输出。想象一下,从几百个随机数组合中,生成梵高的画作、汽车,甚至是人脸。这是多么神奇的事情!

谷歌语言翻译服务、Alexa 语音识别和自动驾驶汽车均采用深度学习算法。其他相关领域也严重依赖深度学习算法,例如语音合成、新药研发及面部识别等。即使在创意领域,如音乐、绘画和写作,也开始被这项技术所颠覆。事实上,深度学习算法在经济上创造出了深刻的转型升级,这可能引发人类所见到的最大变革之一。

由于免费、强大的计算框架和 API(如 Keras 和 TensorFlow)的传播,运行模型的廉价云服务以及数据的便捷可用性,任何人都可以在几小时内在家中运行深度学习模型。这种“平民化”就解释了为什么对深度学习感兴趣的人数呈爆炸性增长,以及在开放格式 Arxiv 和 NIPS 等专业顶级会议上呈现的众多突破的原因。

本书巧妙地通过抽象数学技能探索各种深度学习算法,讲解了计算机视觉、自然语言处理、强化学习和无监督深度学习等深度学习领域的具体商业应用的案例。读者可以通过深入理解业务应用程序,了解有关各个领域未来发展的应用示例。

本书简要介绍了整个深度学习领域的最新算法,其主要目的为使算法更为实

用:解释和说明在几个应用领域中使用的一些重要的深度学习算法,特别是对核心业务有重大影响的深度学习算法。本书面向中高级专业人员以及对机器学习有基本了解的入门级专业人员,以及那些想要了解深度学习并将其用于开发商务应用的人,旨在为从业人员提供实用有效的实施方法。书中过滤掉了令人无所适从的统计学和线性代数推导,为读者提供了如何为商业模式制作简单动手工具的方法和技巧。

本书首先介绍了深度学习架构,并给出了简要历史背景。接下来介绍了深度学习的最先进实例,与传统的机器学习算法相比,其具有更好的应用前景。书中涵盖了推荐系统和自然语言处理的应用,包括能够捕捉语言翻译模型丰富性的递归神经网络(RNN)。最后介绍了深度学习模型在金融风险评估、控制和机器人技术及图像识别中的应用。在书中,你可以了解到其产品中采用该技术的关键公司和初创公司,还可以找到有用的链接以及一些关于如何使用 Keras 和 Python 中的一些实际的代码示例和训练深度学习模型的示例、技巧和见解。

# 目 录

## 第一部分 背景和基础知识

### 第 1 章 绪 论/3

- 1.1 范围和动机/4
- 1.2 深度学习领域的挑战/5
- 1.3 目标受众/5
- 1.4 本书结构/5

### 第 2 章 深度学习概述/7

- 2.1 冬去春来/8
- 2.2 为什么 DL 不同/10
  - 2.2.1 机器时代/12
  - 2.2.2 对 DL 的一些批评/13
- 2.3 资 源/14
  - 2.3.1 图 书/14
  - 2.3.2 简 讯/14
  - 2.3.3 博 客/15
  - 2.3.4 在线视频和课程/16
  - 2.3.5 播 客/16
  - 2.3.6 其他网络资源/17
  - 2.3.7 从一些不错的地方开始学习/18
  - 2.3.8 会 议/19
  - 2.3.9 其他资源/19
  - 2.3.10 DL 框架/20

2.3.11 DL 即服务(DLAS,DL As a Service)/21

2.4 最近的发展/24

2.4.1 2016 年/24

2.4.2 2017 年/24

2.4.3 演化算法/25

2.4.4 创造力/25

### 第 3 章 神经网络模型/27

3.1 神经网络简史/27

3.1.1 多层感知器/29

3.2 什么是神经网络/30

3.3 玻耳兹曼机器/32

3.3.1 受限玻耳兹曼机器/33

3.3.2 深度信念网络/34

3.3.3 深度玻耳兹曼机器/35

3.4 卷积神经网络/36

3.5 深度自动编码器/37

3.6 递归神经网络/37

3.6.1 强化学习的 RNN/40

3.6.2 LSTM/41

3.7 生成模型/43

3.7.1 变分自动编码器/43

3.7.2 生成性对抗网络/45

## 第二部分 深度学习:核心应用

### 第 4 章 图像处理/51

4.1 CNN 图像处理模型/52

4.2 ImageNet 及其他/54

4.3 图像分割/57

4.4 图像标题/58

4.5 视觉问答(VQA)/60

4.6 视频分析/62

- 4.7 GAN 和生成模型/65
- 4.8 其他应用/67
  - 4.8.1 卫星图像/68
- 4.9 新闻和公司/69
- 4.10 第三方工具和 API/70

## 第 5 章 自然语言处理及语音/73

- 5.1 解 析/74
- 5.2 分布式表示/75
- 5.3 知识表示与知识图谱/77
- 5.4 自然语言翻译/81
- 5.5 其他应用/83
- 5.6 多模态学习与问答/85
- 5.7 语音识别/85
- 5.8 新闻与资源/87
- 5.9 总结与思考展望/88

## 第 6 章 强化学习和机器人/89

- 6.1 什么是强化学习/90
- 6.2 传统的 RL/90
- 6.3 DNN 强化学习/92
  - 6.3.1 确定性政策梯度/92
  - 6.3.2 深层确定性政策梯度/93
  - 6.3.3 深度 Q 学习/93
  - 6.3.4 Actor - Critic 算法/95
- 6.4 机器人与控制/97
- 6.5 自动驾驶汽车/99
- 6.6 会话机器人(聊天机器人)/100
- 6.7 新闻聊天机器人/103
- 6.8 应 用/104
- 6.9 展 望/105
- 6.10 自动驾驶汽车的相关新闻/105

## 第三部分 深度学习:商务应用

### 第7章 推荐算法和电子商务/111

- 7.1 在线用户行为/111
- 7.2 重新定向/112
- 7.3 推荐算法/113
  - 7.3.1 协同过滤器/114
  - 7.3.2 RS 的深度学习算法/115
  - 7.3.3 Item2Vec/116
- 7.4 推荐算法的应用/117
- 7.5 未来发展方向/118

### 第8章 游戏和艺术/121

- 8.1 早期的国际象棋/121
- 8.2 从国际象棋到围棋/122
- 8.3 其他游戏和新闻/123
  - 8.3.1 Doom/123
  - 8.3.2 Dota/123
  - 8.3.3 其他应用/124
- 8.4 人造角色/124
- 8.5 艺术中的应用/125
- 8.6 音乐/127
- 8.7 多模态学习/129
- 8.8 其他应用/130

### 第9章 其他应用/135

- 9.1 异常检测与欺诈/135
  - 9.1.1 欺诈预防/137
  - 9.1.2 网上评论的欺诈行为/138
- 9.2 安保及防范/139
- 9.3 预测/140
  - 9.3.1 交易和对冲基金/142

- 9.4 医学和生物医学/143
  - 9.4.1 图像处理和医学图像/144
  - 9.4.2 生物组学/146
  - 9.4.3 药物发现/147
- 9.5 其他应用/148
  - 9.5.1 用户体验/148
  - 9.5.2 大数据/149
- 9.6 未 来/150

## 第四部分 机遇与展望

### 第 10 章 深度学习技术的商务影响/153

- 10.1 深度学习机会/154
- 10.2 计算机视觉/155
- 10.3 AI 助手/155
- 10.4 法 律/157
- 10.5 放射学和医学图像/157
- 10.6 自动驾驶汽车/158
- 10.7 数据中心/159
- 10.8 利用 DL 建立竞争优势/159
- 10.9 人 才/160
- 10.10 光有准确度还不够/161
- 10.11 风 险/162
- 10.12 当个人助理变得比我们更聪明时/162

### 第 11 章 新近研究和未来方向/165

- 11.1 研 究/165
  - 11.1.1 注意机制/166
  - 11.1.2 多模式学习/167
  - 11.1.3 一次性学习/168
  - 11.1.4 强化学习和推理/169
  - 11.1.5 生成神经网络/170
  - 11.1.6 生成性对抗神经网络/171