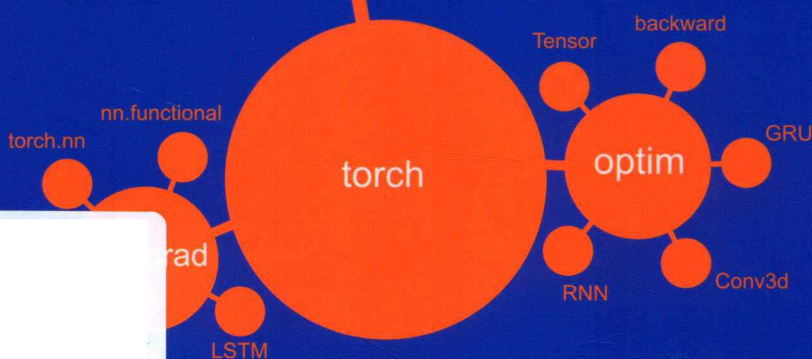


TURING 图灵原创

PyTorch 深度学习原理与实战

集智俱乐部 © 著



中国工信出版集团

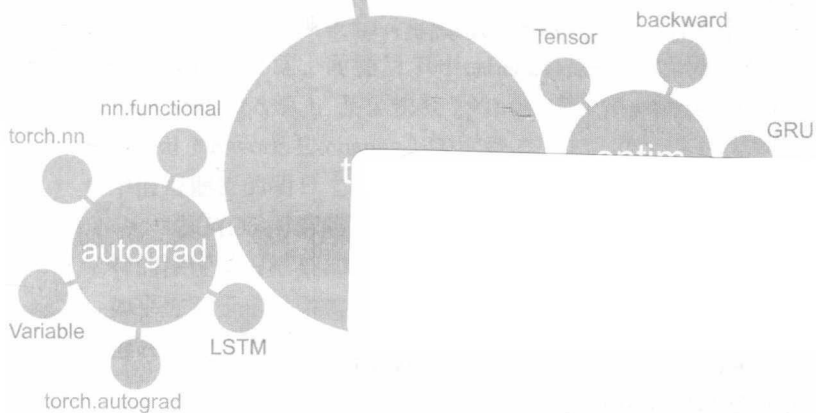
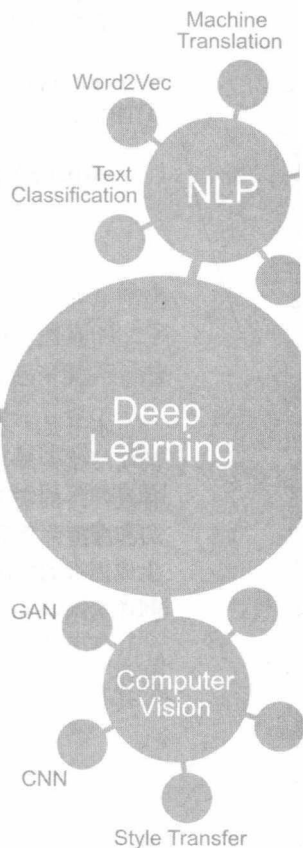


人民邮电出版社
POSTS & TELECOM PRESS

TURING 图灵原创

PyTorch 深度学习原理与实战

集智俱乐部 © 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PyTorch实战 / 集智俱乐部著. --
北京: 人民邮电出版社, 2019.8
(图灵原创)
ISBN 978-7-115-51605-3

I. ①深… II. ①集… III. ①机器学习 IV.
①TP181

中国版本图书馆CIP数据核字(2019)第131055号

内 容 提 要

本书是一本系统介绍深度学习及开源框架 PyTorch 的入门书。全书注重实战, 每章围绕一个有意思的实战案例展开, 不仅循序渐进地讲解了 PyTorch 的基本使用、神经网络的搭建、卷积神经网络和循环神经网络的实现, 而且全面深入地介绍了计算机视觉、自然语言处理、迁移学习, 以及最新的对抗式学习和深度强化学习等前沿技术。读者通过阅读本书, 可以轻松入门深度学习, 学会构造一个图像识别器, 生成逼真的图画, 让机器理解单词与文本, 让机器作曲, 教会机器玩游戏, 还可以实现一个简单的机器翻译系统。

本书适用于人工智能行业的软件工程师、对人工智能感兴趣的学生, 也非常适合作为深度学习培训教程。

-
- ◆ 著 集智俱乐部
责任编辑 张霞
责任印制 周昇亮
 - ◆ 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>
北京市艺辉印刷有限公司印刷
 - ◆ 开本: 800×1000 1/16
印张: 21.5 彩插: 4
字数: 508千字 2019年8月第1版
印数: 1-4000册 2019年8月北京第1次印刷
-

定价: 79.00元

读者服务热线: (010)51095183转600 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

站在巨人的肩上
Standing on Shoulders of Giants



iTuring.cn

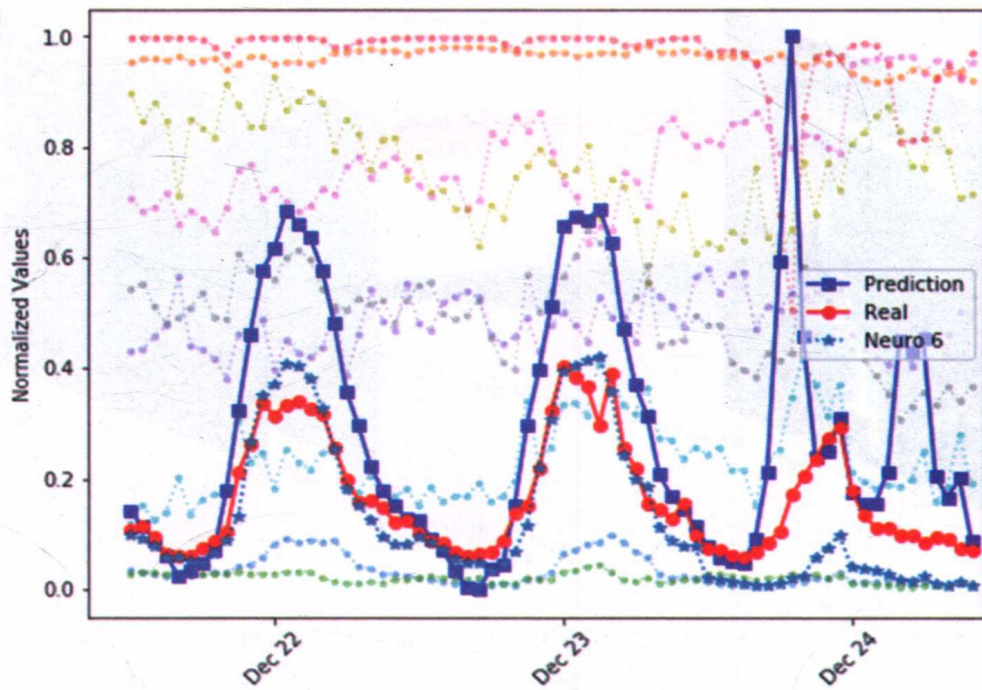


图 3.25 6号神经元激活

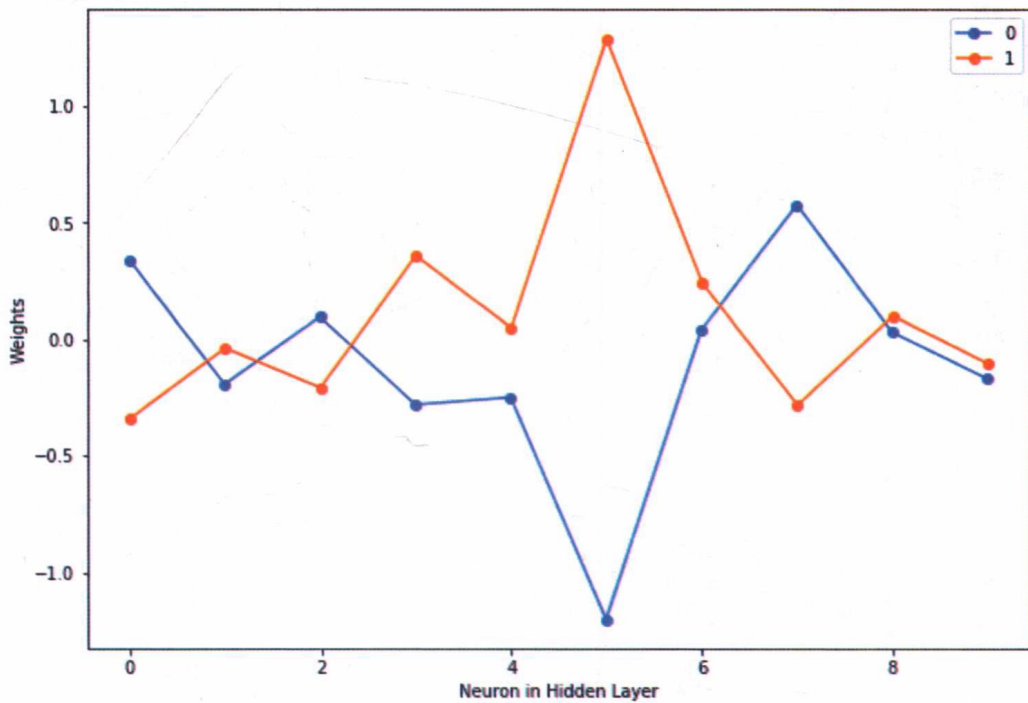
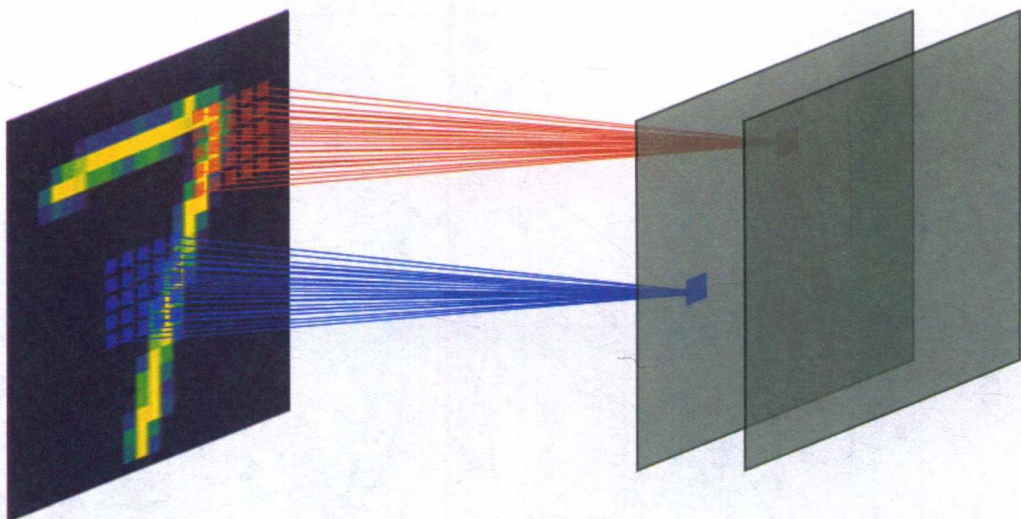


图 4.10 隐含层到输出层两个节点的连边权重



特征图

图 5.15 多个卷积核、特征图和神经网络权重

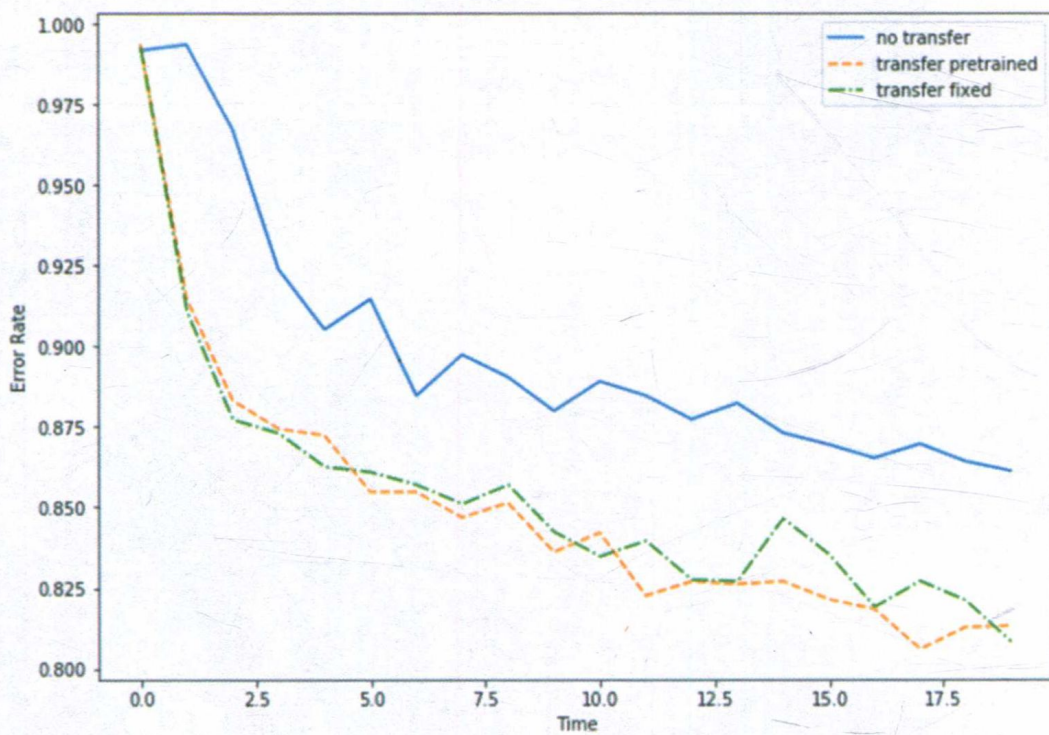


图 6.16 3 种迁移训练准确率结果

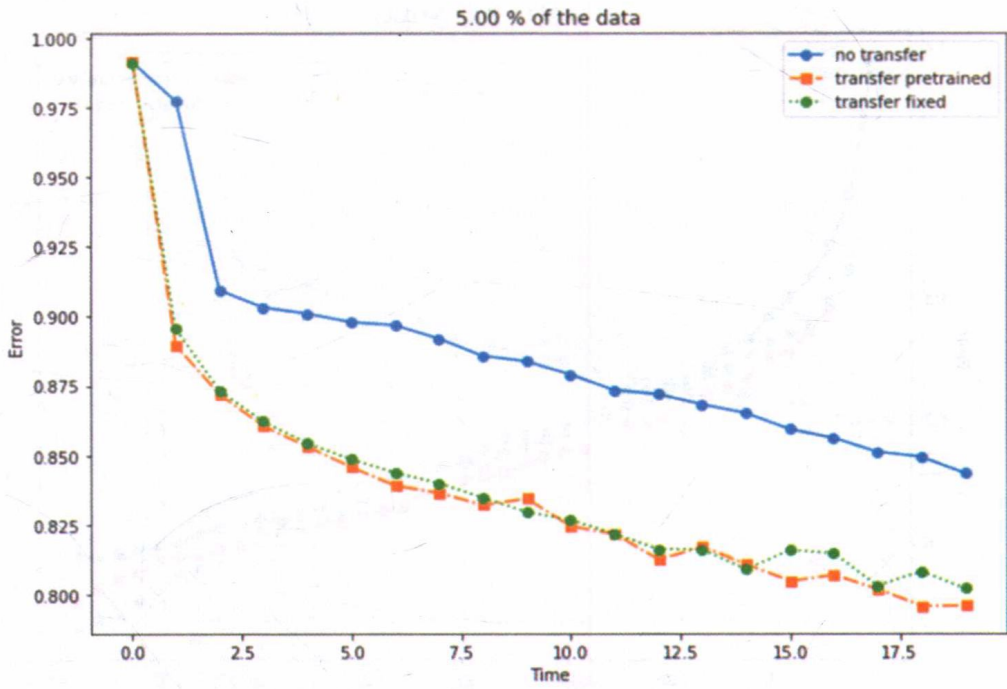


图 6.17(a) 大模型下 3 种迁移方式的准确率曲线 (数据量 5%)

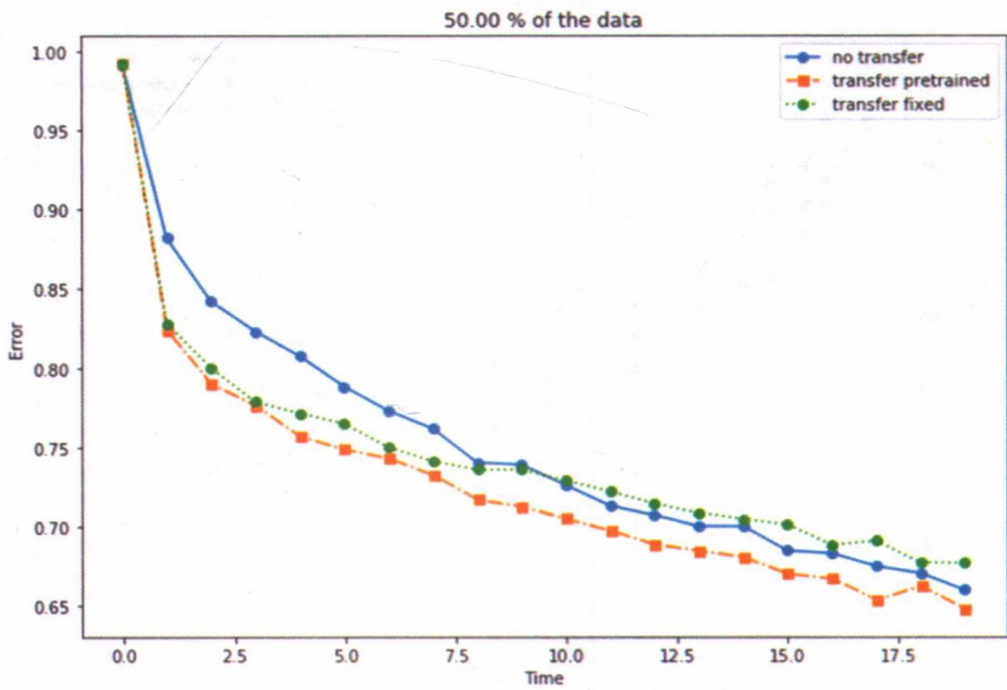


图 6.17(b) 大模型下 3 种迁移方式的准确率曲线 (数据量 50%)

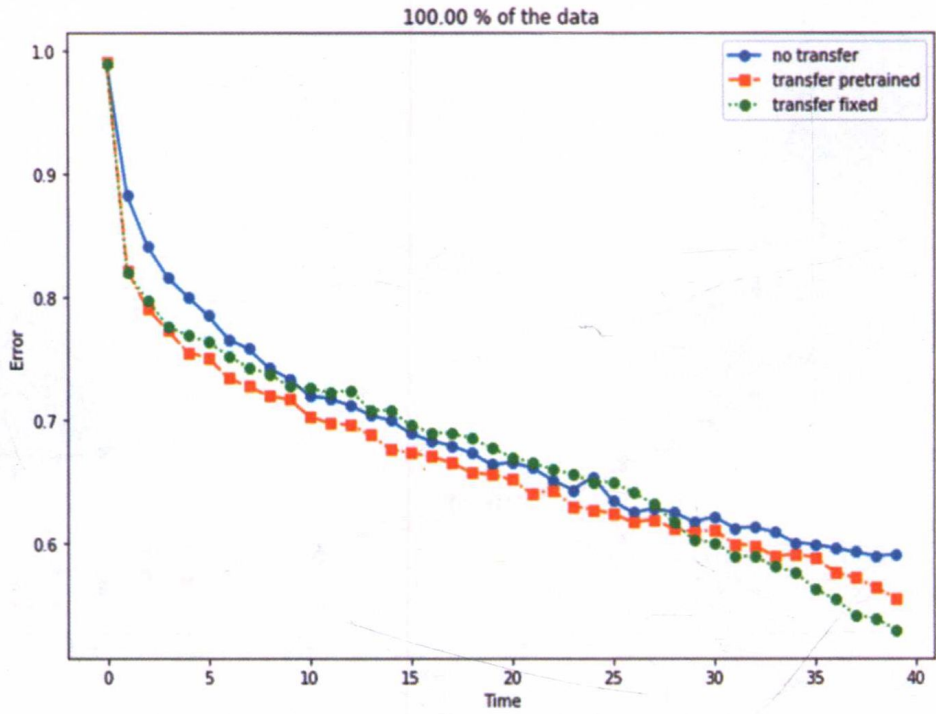


图 6.17(c) 大模型下 3 种迁移方式的准确率曲线 (数据量 100%)

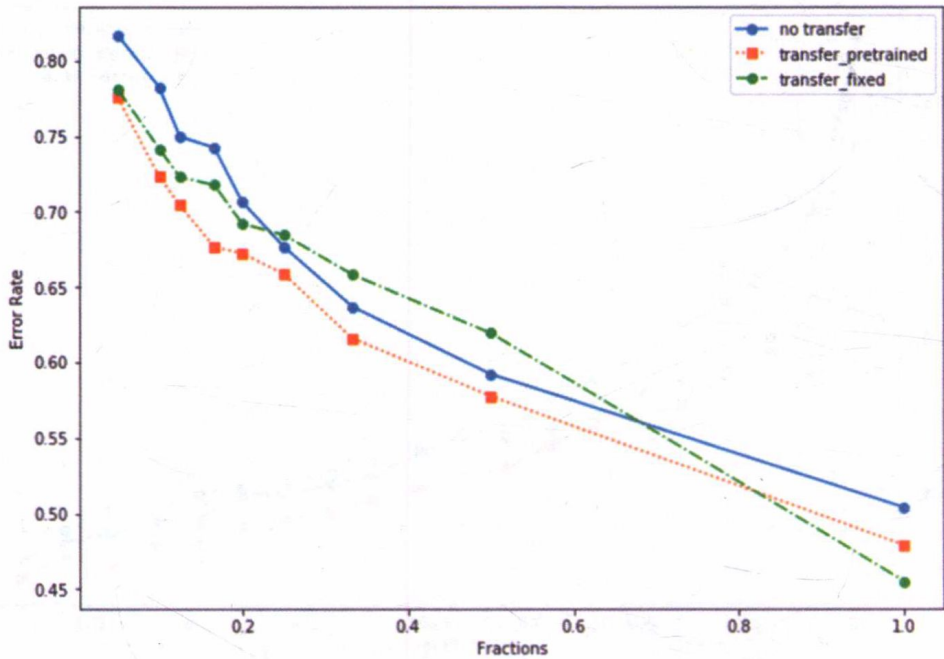


图 6.18 不同数据量下的错误率曲线图

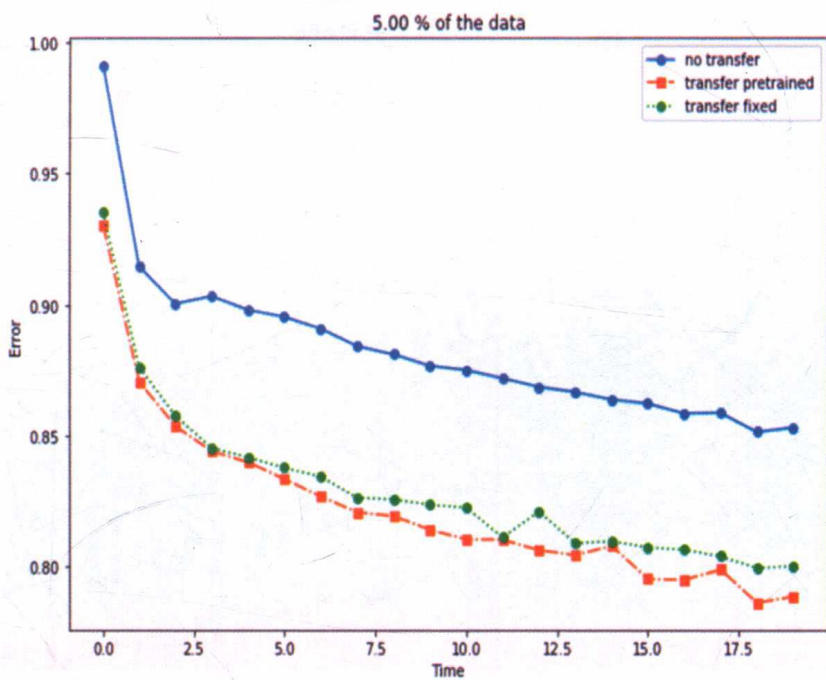


图 6.19(a) 不同数据量对小模型训练结果的影响 (数据量 5%)

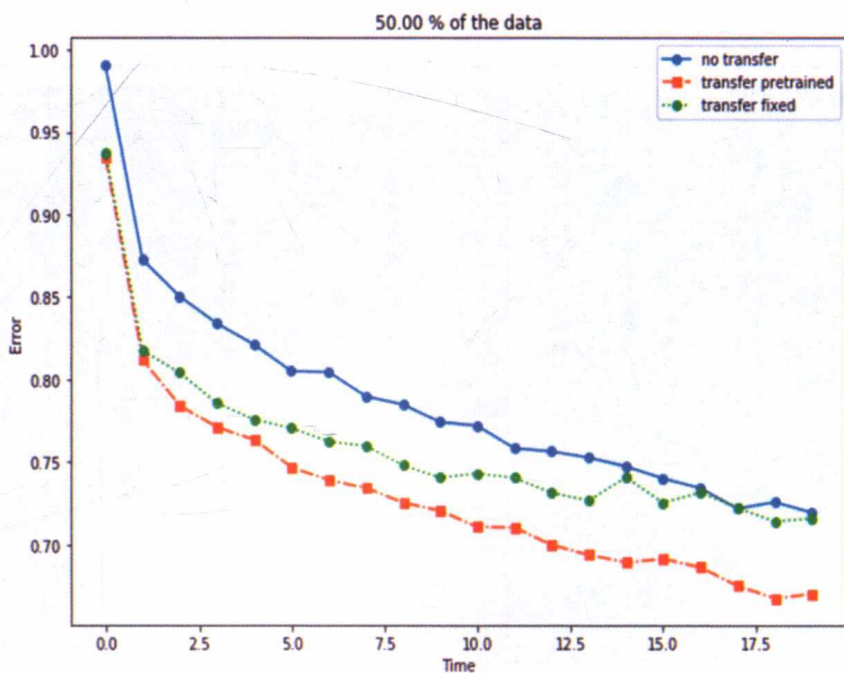


图 6.19(b) 不同数据量对小模型训练结果的影响 (数据量 50%)

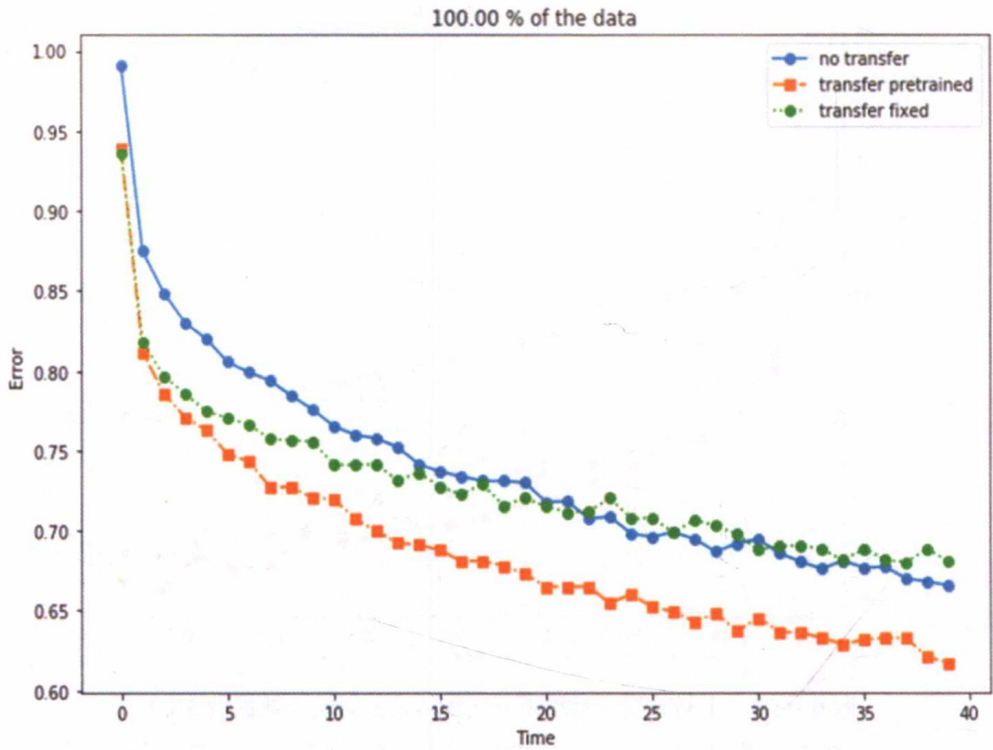


图 6.19(c) 不同数据量对小模型训练结果的影响 (数据量 100%)

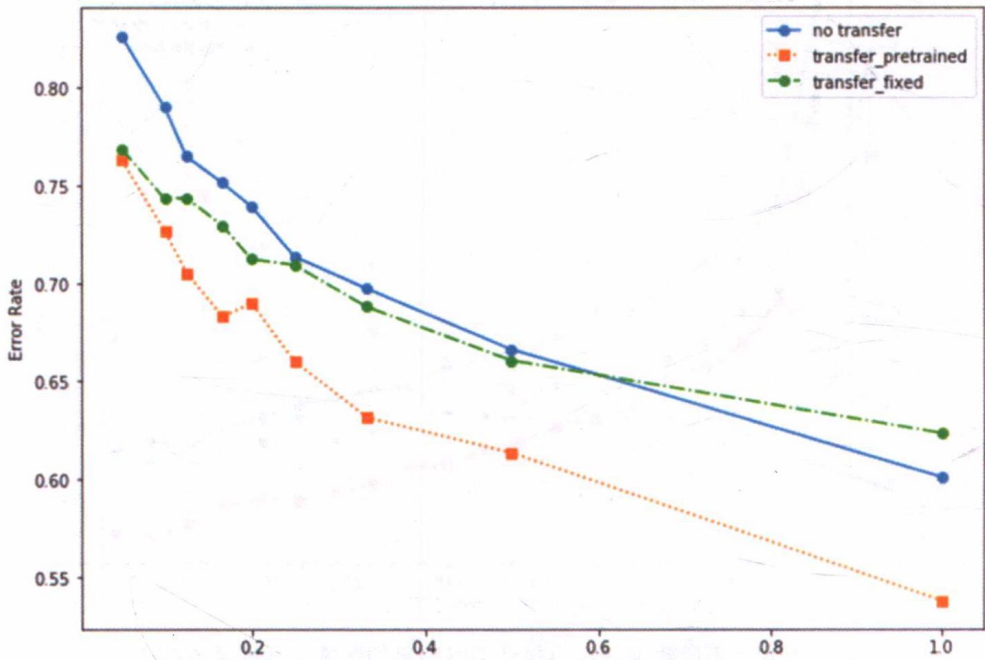


图 6.20 不同数据量下的错误率曲线

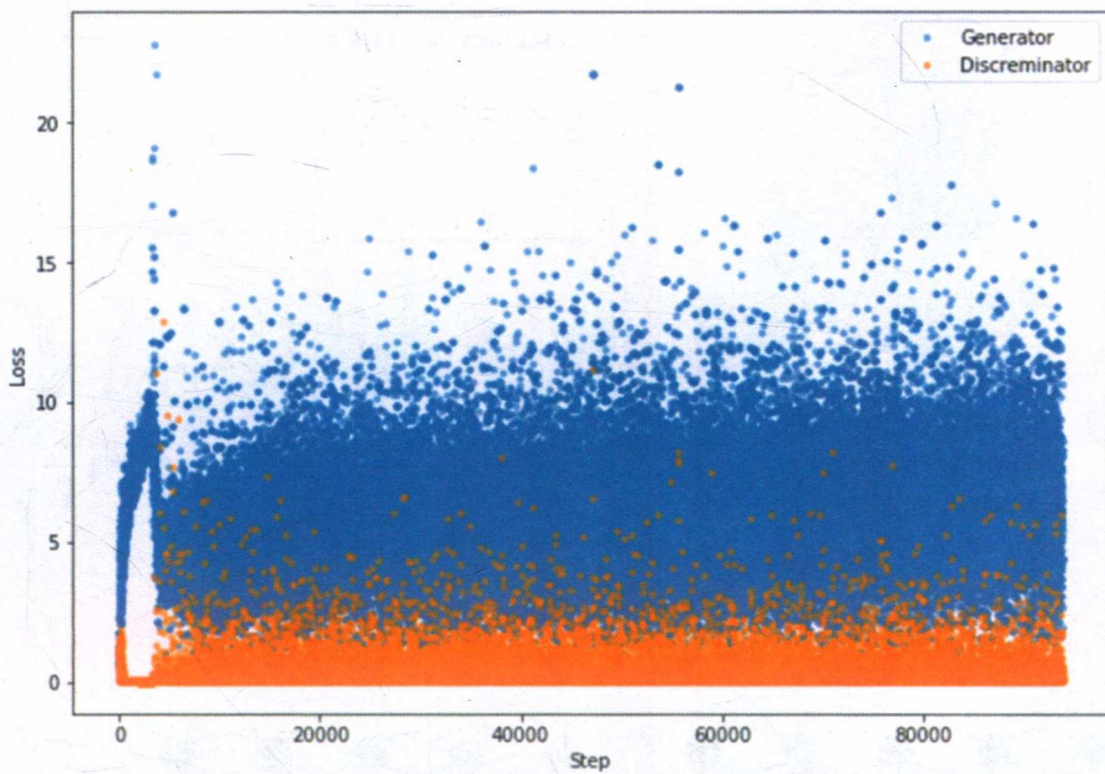
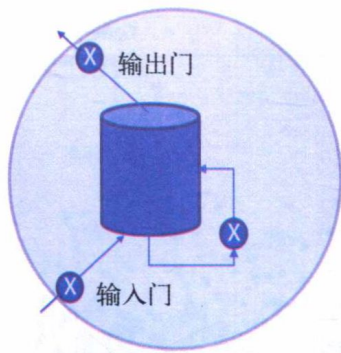


图 8.30 GAN 的学习曲线



图 9.12 利用词向量绘制的“词汇的星空”



$$i_t = \sigma(W_{ii}x_t + W_{hi}h_t + b_i)$$

$$f_t = \sigma(W_{if}x_t + W_{hf}h_t + b_f)$$

$$o_t = \sigma(W_{io}x_t + W_{ho}h_t + b_o)$$

$$g_t = \tanh(W_{ig}x_{t-1} + W_{hg}h_{t-1} + b_g)$$

$$c_t = f_t * c_{t-1} + i_{t-1} * g_{t-1}$$

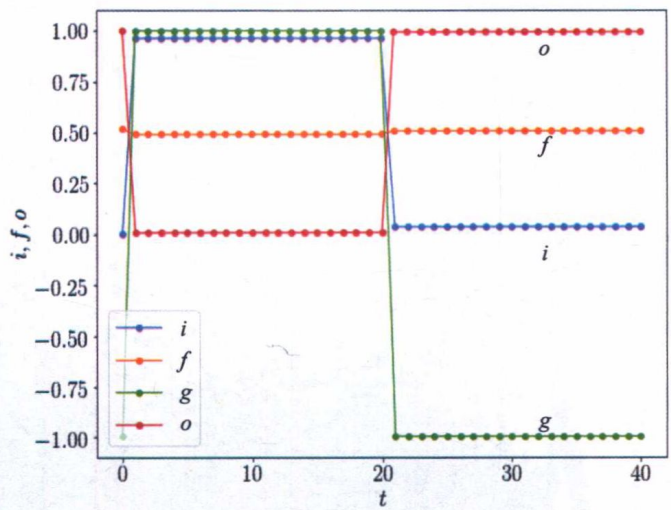


图 10.16 LSTM 单元内部的 3 个门的开启情况

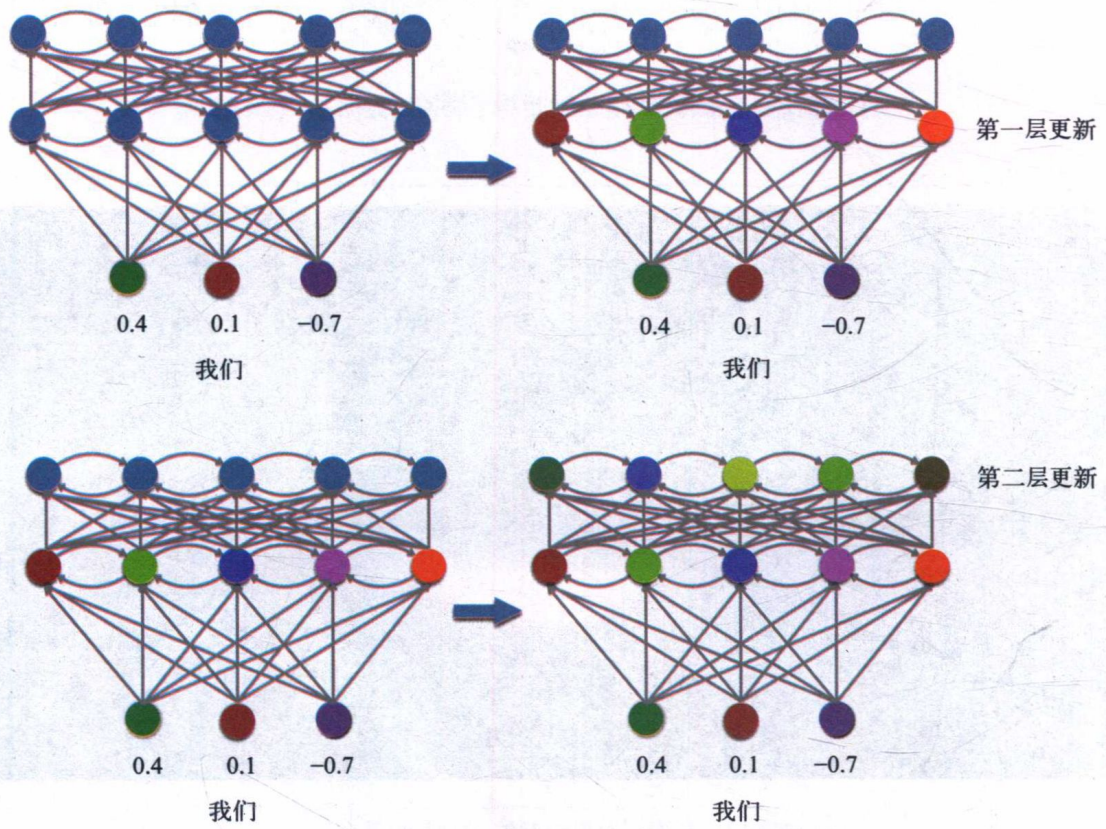


图 11.7 循环神经网络状态的逐层传播

推 荐 序

“理解复杂世界”是生活在这个星球的所有人类共同的愿望。

4000 多年前，埃及人通过观察发现，当夏天黎明天狼星于东方升起时，尼罗河就会开始洪水泛滥。中国的先民由于积累了大量的观察记录，发展出了用来指代理性思考于脑中仿真的文字符号——“预”。“预”字的左边（予）表示“通过”，右边（页）表示“头脑”，合起来是指让想象情景过脑，这也象征了人类独特的智能行为——“预测”。

随着智能的逐步累积，人类观察世界的工具也有了质的飞越，从最早的肉眼观察记录，逐步发展出了通过实验和理论来推演思辩的方式。但随着知识的累积，人类发现世界上难以理解的事情与日俱增，这时才意识到过去犯了一个严重的错误，那就是企图用“简化的模型来描述世界”。问题是，当这个世界不简单的时候，任何企图化繁为简的举动都会让我们离真实越来越远。

既然简化模型这条路不可行，那就直接正面挑战这个世界的复杂性吧！深度学习开山鼻祖辛顿为我们带来了挑战复杂性的最新武器。从语音识别、机器视觉，到自然语言，这些我们过去认为机器不可能做到的任务，都被深度学习逐步破解。深度学习的关键正是计算机不再只靠人为的规则编程，而是用数据进行编程，重现这个世界的复杂性。

但现实中的深度学习工具并不友好。例如，TensorFlow 虽然是一款主流的深度学习框架，但其语法难以跨版本兼容，函数改名频率过高，语法啰唆繁杂。虽然有了挑战复杂度的关键神器，但是要想将它切实应用，还要有让人思维更清晰的分析框架。

2017 年初，Facebook 公司推出了全新的深度学习框架——PyTorch。

在深度学习顶级会议 ICLR 的提交论文中，提及 PyTorch 的论文数从 2017 年的 87 篇激增到 2018 年的 252 篇，而提及 TensorFlow 的论文数量却没有太大的起伏（从 2017 年的 228 篇提升至 2018 年的 266 篇），甚至快被 PyTorch 追平了。同时，随着 PyTorch 1.0 的问世以及 ONNX（Open Neural Network Exchange）深度学习开发生态的逐渐完备，PyTorch 无疑成为众多深度学习框架中值得期待的明日之星。

我们在学习深度学习时常常会有一个疑问：既然深度学习要正面解决世界的复杂性，为何现有的深度学习框架却处处是人为的简化呢？比如图片必须固定大小，句子必须固定长度，等等。如果你对于这种束缚感到厌倦，那么使用动态计算图的 PyTorch 可能会是最好的选择。如果你又担心动态计算图难以理解，那么本书将会是你在理解 PyTorch 过程中的最佳帮手，它将神经网络、计算图、自动微分、梯度反传等概念用清晰的文字表达了出来。更重要的是，它很少用到数学公式。

学习技术的过程其实和深度学习的模型一样，需要通过梯度下降的引导，才能让复杂的问题逐渐走向最优解。任何新事物的学习，只有遵循正确的学习路线，才有可能将基础打牢，融会贯通。尤其是对于深度学习这样处于发展中的年轻学科，从各种杂乱的来源中找出接近真相的信息，恐怕是学习上最大的障碍。

由张江老师领衔的集智俱乐部一直是国内复杂性系统与深度学习社群中的领头羊，长期为推广技术和培育新生科研种子而努力。集智俱乐部创作的这本《深度学习原理与 PyTorch 实战》通过丰富的案例和清晰的讲解，带你找到良好的深度学习修炼路线，直达获得最佳学习状态，而不必像随机梯度下降般迂回绕路。

读完本书，你会发现，有了强大的工具和便捷的方法，深度学习竟然可以如此简单。

尹相志

台湾微软数据科学金牌讲师

中国微软加速器专家顾问，集智学园人工智能金牌讲师

作者简介

集智俱乐部 (Swarma Club)，成立于 2003 年，是一个从事学术研究、享受科学乐趣的探索者团体，也是国内最早研究人工智能、复杂系统的科学社区，倡导以平等开放的态度、科学实证的精神，进行跨学科的研究与交流，力图搭建一个中国的“没有围墙的研究所”。目前已出版著作有《科学的极致：漫谈人工智能》和《走近 2050：注意力、互联网与人工智能》，译作有《深度思考：人工智能的终点与人类创造力的起点》。

编写

- 张江 集智俱乐部创始人、集智学园创始人、北京师范大学系统科学学院教授、腾讯研究院特聘顾问
- 张章 北京师范大学系统科学学院科研助理
- 朱瑞鹤 机器视觉算法工程师
- 胡胜 中国地质大学（武汉）博士研究生，研究领域：测绘科学与技术
- 胡乔 工学硕士、结构工程师、算法爱好者
- 王硕 东北大学模式识别与智能系统硕士、彩云科技算法工程师，研究领域：深度学习、雾霾预报、复杂网络嵌入
- 周金阳 果壳网算法工程师
- 苏尚君 文因互联软件工程师、集智学园 PyTorch 课程杰出贡献助教，曾任 Udacity 机器学习助教
- 王婷 北京集智会自然科学研究中心首席秘书长
- 姜晓芳 北京航空航天大学博士研究生、特许金融分析师（CFA）持证人、金融信息化研究所资深研究员
- 丛恺 香港中文大学计算机科学在读
- 胡鹏博 生物物理所神经科学硕士在读

审稿

- 刘福洋 南丹麦大学应用物理专业、Tradeshift 机器学习软件工程师
- 胡新兴 虚拟现实产品经理

统筹

张 倩 集智学园联合创始人兼 CEO，毕业于南京信息工程大学信息与控制学院

李周园 清华大学博士，集智学园 PyTorch 课程杰出贡献助教

刘培源 北京集智会自然科学研究中心首席媒体运营官

封面设计

王建男 集智学园首席商务官

前 言

在 21 世纪的第二个十年里，科技界最大的进展恐怕非人工智能莫属了。无论是战胜人类围棋高手的 AlphaGo，还是遍布全国车站的人脸识别系统，配备了深度学习技术的最新人工智能展现了它无限大的势能，并已经进入到我们的日常生活中。

人工智能（Artificial Intelligence，简称 AI），顾名思义，就是通过计算的方式模拟、延伸和扩展人的智能。它作为计算机科学的一个分支，早在 1956 年就诞生了。然而，长久以来，人工智能的发展却不能与它的名字相匹配。尽管早期的人工智能在数学定理证明、推理、棋类游戏上取得了长足的进步，但是在拟人化的形象思维方面却与人类相差甚远。例如，一个两岁的小孩能清楚地认出爸爸和妈妈，但是人工智能却不能。

不过近年来，人工智能的发展正在试图摆脱人们对它的刻板印象。采用深度神经网络技术的人工智能同样可以非常好地进行“形象化”思维。例如，现在人工智能的人脸识别准确度已经达到了 99.7%，超过了人类的识别准确度 97.3%。2017 年 1 月，百度大脑的人工智能程序参加了《最强大脑》节目，在人脸识别和声纹识别上挑战了人类顶尖高手，并最终完胜了人类的“最强大脑”。如今，人工智能已经可以在自己的弱项上战胜人类了。

在 2016 年 3 月和 2017 年 5 月，AlphaGo 分别与世界围棋冠军李世石和柯洁进行了举世瞩目的比赛。可以看到，配备了深度强化学习技术的人工智能可以像人类围棋高手那样具有出色的大局观，甚至具有一定的创造力。这种表现是单纯依靠逻辑推理和搜索的人工智能远远无法达到的。更有甚者，DeepMind 团队在 2017 年 10 月发表在《自然》上的文章提出了一个新版本的 AlphaGo——AlphaGo Zero，它完全凭借自己的“左右互搏”，而无须任何人类经验，就可以达到世界围棋的顶尖水平，远远超越人类。

然而，这些有关人工智能的新闻会给我们造成一种错觉：人工智能是一种高科技，只有谷歌、微软等大公司才有可能应用，与我们普通人或者小公司毫无关系。事实并非如此，随着各大公司开源了他们的深度学习框架和平台，我们每一个普通企业或者个人都可以快速地应用人工智能技术。你只要有一台笔记本电脑，就可以轻松玩转深度学习，实现诸如人脸识别、图像生成、机器翻译、聊天机器人等强大的人工智能功能。

其实，人工智能早已渗透到了我们的身边。例如，当我们使用导航系统播报路况的时候，林志玲或者郭德纲的声音就会从手机或汽车音响里播放出来。难道林志玲或郭德纲会把成千上万种可能的路况信息都念一遍吗？答案显然是否定的，这是运用了人工智能中的语音合成技术。

有一款 App 叫作 Prisma。你只要上传自己的照片，再选择一张风格图片（例如莫奈的画作），