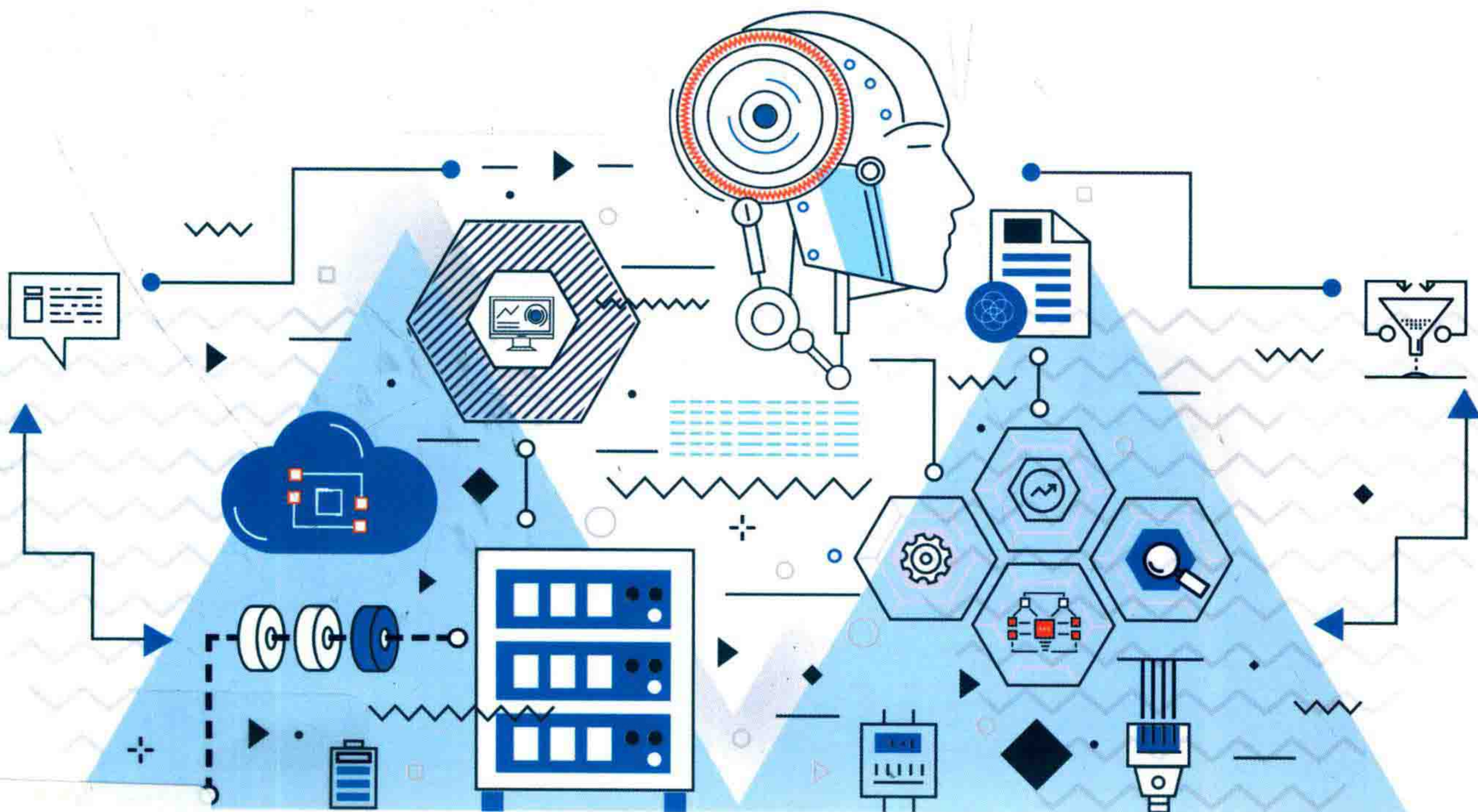


機械学習と深層学習 —C言語によるシミュレーション

机器学习与深度学习

通过C语言模拟

[日] 小高知宏 著
申富饶 于德 译



機械学習と深層学習 —C言語によるシミュレーション

机器学习与深度学习

通过C语言模拟

[日] 小高知宏 著

申富饶 于德 译



图书在版编目 (CIP) 数据

机器学习与深度学习：通过 C 语言模拟 / (日) 小高知宏著；申富饶，于德译．—北京：机械工业出版社，2018.5

(智能系统与技术丛书)

ISBN 978-7-111-59994-4

I. 机… II. ①小… ②申… ③于… III. 机器学习—研究 IV. TP181

中国版本图书馆 CIP 数据核字 (2018) 第 093230 号

本书版权登记号：图字 01-2018-0804

Original Japanese language edition

KIKAI GAKUSHU TO SHINSOU GAKUSHU

— C GENGO NI YORU SIMULATION —

by Tomohiro Odaka

Copyright © 2016 Tomohiro Odaka

Published by Ohmsha, Ltd.

Chinese translation rights in simplified characters arranged with Ohmsha, Ltd.

through Japan UNI Agency, Inc., Tokyo

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from the publisher.

本书中文简体字版由 Ohmsha, Ltd. 通过 Japan UNI Agency, Inc. 授权机械工业出版社独家出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

机器学习与深度学习：通过 C 语言模拟

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：张梦玲

责任校对：殷虹

印刷：北京市兆成印刷有限责任公司

版次：2018 年 6 月第 1 版第 1 次印刷

开本：186mm × 240mm 1/16

印张：12.25

书号：ISBN 978-7-111-59994-4

定价：59.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88379426 88361066

投稿热线：(010) 88379604

购书热线：(010) 68326294 88379649 68995259

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

P R E F A C E

前 言

当前，关于人工智能的研究非常引人瞩目。作为其支柱之一的深度学习，是人工智能研究至今为止所积累的机器学习成果。特别是在语音识别、图像识别、行为知识获取等方面，深度学习取得了非常大的成功。

本书以非常容易理解的方式解说了人工智能研究中机器学习的各领域知识，以这些知识为前提，说明了深度学习是什么。本书不是单纯地罗列概念，而是通过适当介绍具体处理流程和程序示例，来具体化地、通俗易懂地介绍这些技术到底是什么。

本书通过例题程序来具体说明机器学习和深度学习的相关技术。一般而言，运行深度学习程序的机器必须有强大的计算能力才行。但是，在本书的例题程序中，经过精心设计，将深度学习处理的核心骨架部分提取了出来，在普通的个人计算机上也能运行。假设运行环境为 Windows，通过 Visual Studio 等开发工具来编译程序，给定适当的数据，就能够执行例题程序。通过尝试执行程序，读者能够更具体、更深入地理解机器学习和深度学习相关的技术。

本书是 Ohmsha 先前出版的《はじめての機械学習》的姐妹篇。《はじめての機械学習》广泛叙述了机器学习相关的话题。与此相对的是，本书以深度学习为主题，并对相关内容做了有针对性的处理。本书和上一本书有一些话题是重复的，但本书的说明方式和例题设计，是为了更好地理解深度学习而原创的。

能够完成本书，离不开作者在福井大学从事教育研究活动所积累的经验，感谢福井大学的教职员和学生。此外，感谢 Ohmsha 的同仁给予了出版本书的机会。最后，感谢支持我完成本书的家庭成员（洋子、研太郎、桃子、优）。

2016年4月

目 录

前言	
第 1 章 机器学习 1	
1.1 什么是机器学习..... 1	
1.1.1 深度学习的成果..... 1	
1.1.2 学习、机器学习和深度学习..... 6	
1.1.3 机器学习的分类..... 9	
1.1.4 直至深度学习的机器学习历史..... 15	
1.2 关于本书例题程序的执行环境..... 25	
1.2.1 程序执行的流程..... 25	
1.2.2 程序执行的实际情况..... 27	
第 2 章 机器学习基础 31	
2.1 归纳学习..... 31	
2.1.1 演绎学习和归纳学习..... 31	
2.1.2 归纳学习的例题——股票价格的预测..... 32	
2.1.3 基于归纳学习的股价预测程序..... 37	
2.2 强化学习..... 46	
2.2.1 什么是强化学习..... 46	
2.2.2 Q 学习——强化学习的具体方法..... 48	
2.2.3 强化学习的例题——走迷宫知识的学习..... 53	
2.2.4 强化学习程序的实现..... 56	
第 3 章 群体智能与演化方法 65	
3.1 群体智能..... 65	
3.1.1 粒子群最优化方法..... 65	
3.1.2 蚁群最优化方法..... 67	
3.1.3 蚁群最优化方法的实现..... 70	
3.2 演化方法..... 81	
3.2.1 什么是演化方法..... 81	
3.2.2 基于遗传算法的知识获取..... 84	
第 4 章 神经网络 101	
4.1 神经网络基础..... 101	
4.1.1 人工神经元模型..... 101	
4.1.2 神经网络与学习..... 105	

4.1.3	神经网络的种类	107	5.1.2	卷积神经网络	142
4.1.4	人工神经元的计算方法	108	5.1.3	自编码器的学习方法	145
4.1.5	神经网络的计算方法	115	5.2	深度学习的实现	147
4.2	基于反向传播的神经网络的学习	121	5.2.1	卷积运算的实现	148
4.2.1	感知机的学习过程	121	5.2.2	卷积神经网络的实现	156
4.2.2	反向传播的处理过程	123	5.2.3	自编码器的实现	170
4.2.3	反向传播的实现	125			
第 5 章	深度学习	139	附录 A	生成行李的重量和价值的程序	183
5.1	什么是深度学习	139	附录 B	通过全搜索求解背包问题的程序	185
5.1.1	传统神经网络的局限和深度学习的思路	139	参考文献		189

机器学习

本章介绍什么是机器学习，并对深度学习这一机器学习方法进行说明。首先，介绍近年来被广为关注的深度学习的成果，说明深度学习技术为什么受到人们的重视。然后，阐述什么是学习、机器学习和深度学习，概述目前为止的机器学习研究历史。最后，说明如何执行本书的例题程序。

1.1 什么是机器学习

在本节中，我们会基于多个研究实例来概述深度学习到底能干什么（关于本节所介绍的深度学习系统的具体实现技术，会在第 5 章做进一步说明）。

1.1.1 深度学习的成果

近年来，深度学习（deep learning）技术得到了广泛关注。深度学习被广为关注的原因在于，利用深度学习方法能够完成传统机器学习系统所不能完成的多种情形下的智能处理任务。

表 1.1 给出了基于深度学习所实现的智能处理系统示例。

表 1.1 基于深度学习实现的智能处理系统示例

序号	系 统 名	相 关 论 文	说 明
1	DQN (Deep Q-Network)	Volodymyr Mnih 等: Human-level control through deep reinforcement learning, Nature, Vol. 518, pp. 529-533 (2015)	系统输入游戏画面, 输出控制游戏的操作信号。该系统给出了通过深度学习获得能得到高分的控制方法的例子。
2	ConvNet VGG	Karen Simonyan, Andrew Zisserman: VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION, ICLR 2015 (2015)	用深度学习方法中的一种——卷积神经网络 (Convolutional Neural Network, CNN) 来进行图像识别, 对于标准样例实现了其他方法不可能实现的高水平识别性能
3	CD-DNN-HMM	Frank Seide, Gang Li, Dong Yu: Conversational Speech Transcription Using Context-Dependent Deep Neural Networks, INTERSPEECH 2011, pp. 437-440(2011).	用深度学习进行语音识别的早期研究, 说明了深度学习对于语音识别的有效性

在表 1.1 的第 1 行所给出的 DQN (Deep Q-Network) 中, 采用深度学习的学习系统, 在有些场合发挥了超人类的能力, DQN 就给出了这样的研究实例。

在作为该研究对象的控制系统中, 输入的是有古典风格的电视游戏 (视频游戏) 的游戏画面。系统读入这样的游戏画面后, 根据画面输出对游戏操控杆的操作。也就是说, 这一控制系统实际上是能够玩游戏的计算机玩家。这些游戏包括弹球、ブロック崩し (消消球)、乒乓球以及老粉丝们所怀念的“クレイジークライマー (疯狂登山者)[⊖]”“ロードランナー (奔跑者)[⊖]”等, 都是以前著名的游戏。在玩这些游戏时, 可以根据画面的状态操作操控杆。控制系统会输入画面, 输出操控杆的上、下、左、右以及是否按压按钮等控制信号。

设计这样的控制系统特别困难。目前为止的人工智能研究中所提出的机器学习方法, 要构建这样的控制系统并使之和人类一样进行学习是非常困难的。事实上, 到目前为止, 还没有人报告过这样的研究成果。

针对这一问题, 在 DQN 的研究中, 采用最新的深度学习方法来构建控制系统。

⊖ 1980 年由日本物产发售, 纵向滚屏的动作游戏。

⊖ 1983 年由ブローダーバンド公司发售, 动作拼图游戏。

DQN 将游戏的得分作为线索，系统能自动学会如何更好地操作操控杆。这一过程是完全自动的，无须人类做任何的操作调整。在这个意义下，可以认为 DQN 能根据所看到的画面学习操控杆的操作，执行了与人类所进行的学习同样的处理（图 1.1）。

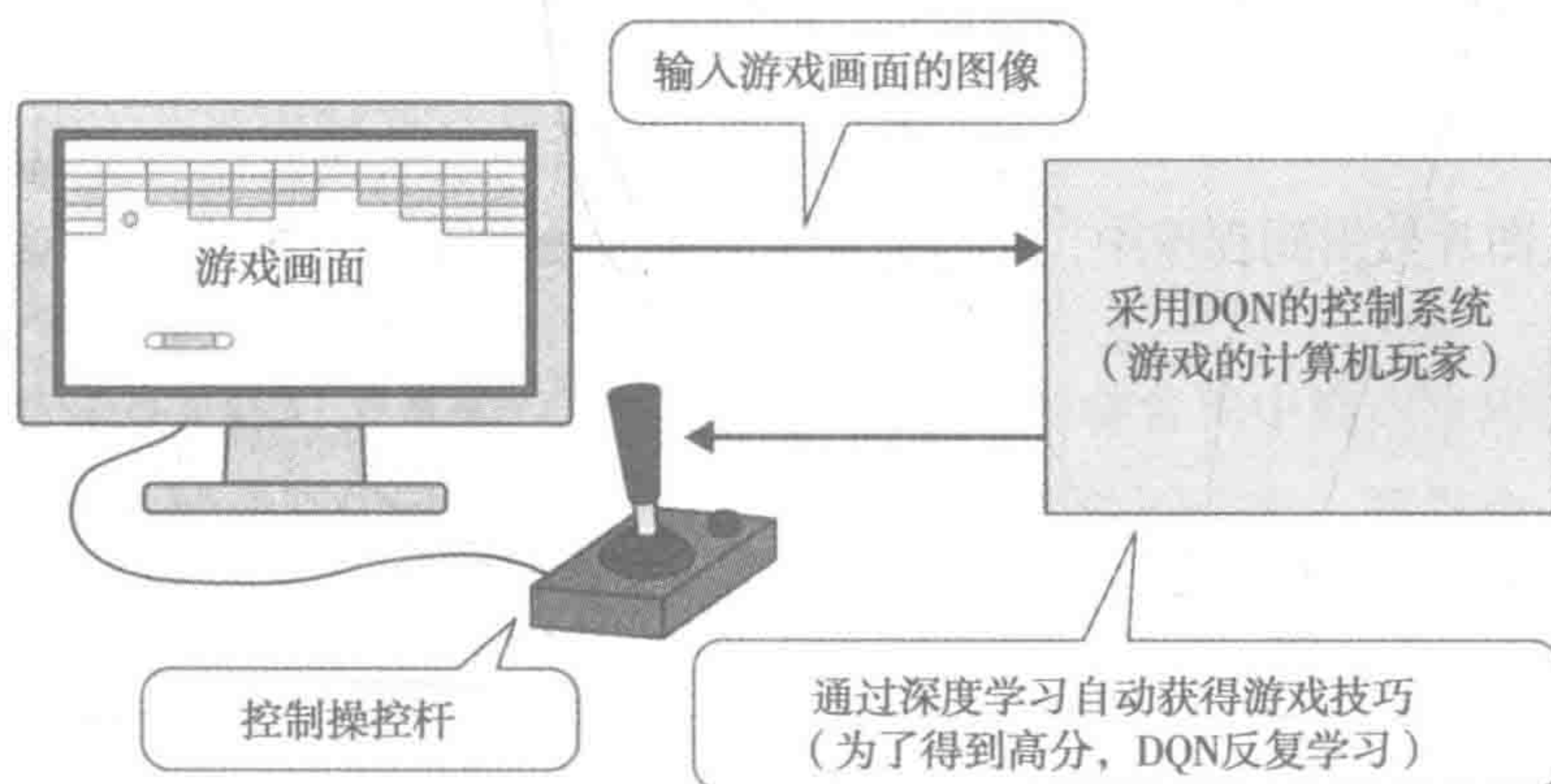


图 1.1 基于 DQN 学习电视游戏

在 DQN 的学习结果中，控制系统在多个游戏上都比人类强，成功学会了操控杆的操作方法。根据论文报告的结果，DQN 所给出的最好性能是“Video Pinball”这一游戏。表 1.1 所给出的论文题目虽然是“人类水平的控制 (human-level control)”，但对于某些游戏却获得了超过人类的能力。

就这样，作为深度学习众多例子之一的 DQN，通过观看画面而学习游戏的玩法，实现了像人类一样的智能处理。而且，其学习结果更是获得了超过人类能力的游戏技巧。这一例子可以启发如下思考：深度学习能够像“人类一样”学习并具有超过人类的潜能。

DQN 所使用的技术中，将传统机器学习中最主要的分支之一——强化学习 (reinforcement learning) 技术和深度学习技术结合起来了。在所使用的深度学习技术中，采用的是学习的核心技术——卷积神经网络（关于这些技术的详细情况，在后面会加以说明）。

在表 1.1 的第二个例子中，给出了将深度学习用于图像识别的研究成果示例。这一研究中使用卷积神经网络来构建用于图像识别的深度学习系统，这一系统能判别输入照

片中所拍的对象是什么。

作为研究对象的例题是称为“The ImageNet Large Scale Visual Recognition Challenge (ILSVRC)”的通过机器学习进行图像处理国际学术竞赛所提供的照片图像数据。这一图像数据由大量的照片组成，照片中有老虎、狮子等动物，有汽车、飞机和坦克等交通工具，还有红酒及蘑菇等各色各样的东西。这一研究的目标是构建一个图像识别系统，该系统能读入图片数据到程序中，然后自动地将图片分类到1000个类别中。

像这样的识别图像中所含物体的课题，对于人类来说是比较容易的，但对于计算机软件而言是一个难题。这一例题所含的图像，既有比较容易识别的图像，也有即使对于人类而言也会一看就产生“这是什么”感觉的、很难识别的图像。在人工智能领域人们提出了多种不同的机器学习方法，但对于识别图像并将之分类的问题而言，这些方法都非常困难。

对此，近年来，使用深度学习技术，在图像分类上达到了过去从未达到的识别精度。表1.1的第二个例子也是其中之一，卷积神经网络的处理结果超出了传统机器学习方法的界限。在这一例子中，输入的是长宽均为224像素的RGB图像，输出的是该图像属于1000个类中哪一个类的信息。这个例子和先前的DQN例子一样，将原始的图像不做任何变换输入到系统中，通过深度学习系统来学习如何进行图像的判别(图1.2)。

在上述学习过程中用到了包括图片及其所属正确类别的数据集。对于输入数据预先给定正确的解，通过分析这些数据经过学习产生一种对于输入数据与正解之间特定的映射关系，其可以映射出新的实例。这样的学习称为监督学习[⊖](supervised learning)。已经知道正确解的数据集合称为学习数据集或者训练数据集(training dataset)。在这一例子中，采用由大量学习数据所构成的学习数据集来自动学习如何进行图像的识别。

在学习结束后，给系统输入测试用的数据，来检查系统在多大程度上得到了正确的

⊖ 从日语直译是“有教师学习”，似乎比中文的一般译法“监督学习”更形象，本书按中文习惯采用“监督学习”的译法。——译者注

解。这样的数据集称为检查数据集或测试数据集 (test dataset)。在这一研究中,将深度学习技术所学习的结果和其他学习方法所学习的结果在测试数据集上进行了比较,证明了深度学习所得到的识别精度比其他方法高。

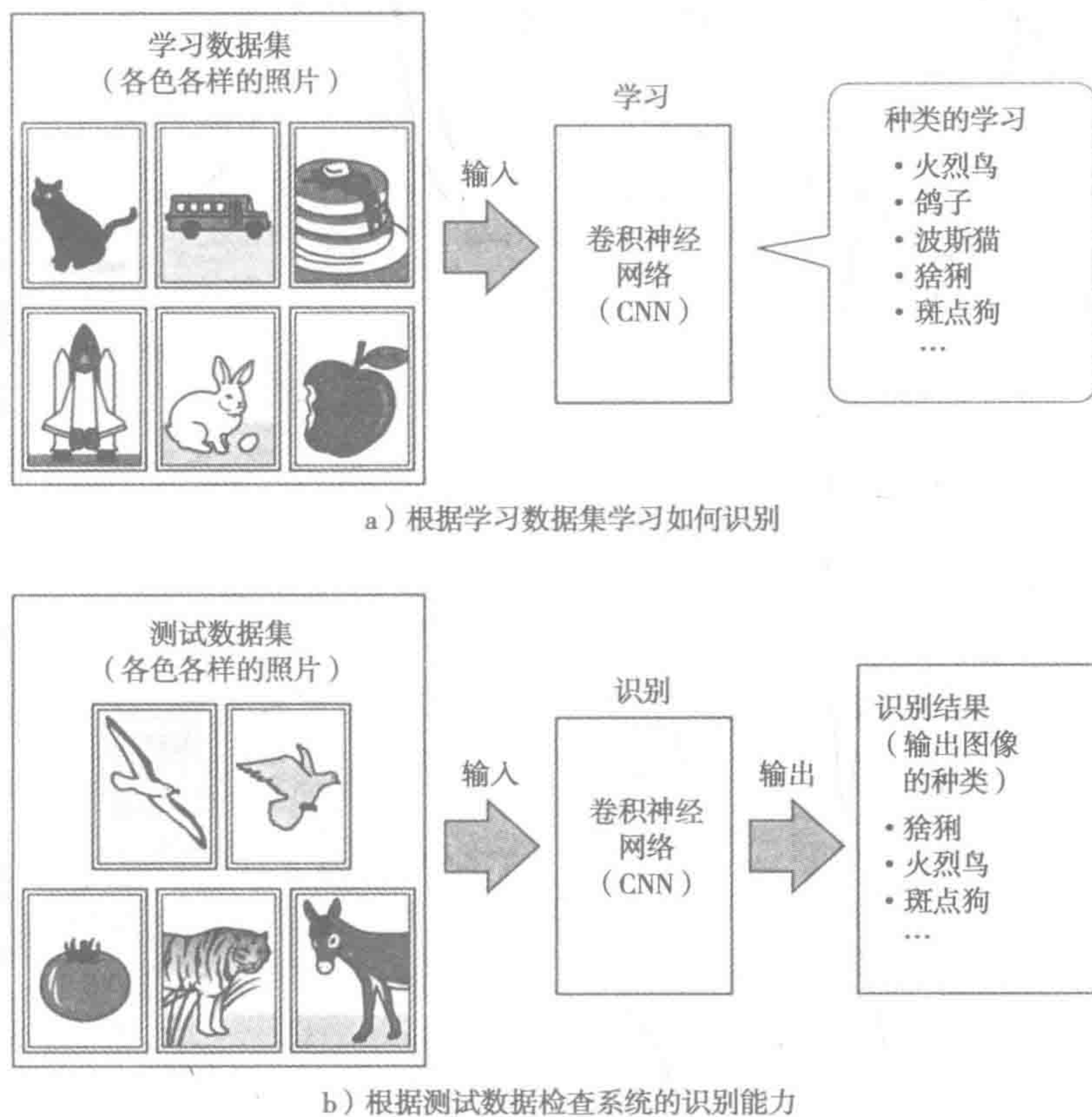


图 1.2 基于卷积神经网络的图像识别

将深度学习应用于图像识别的研究非常活跃,除上述研究之外,研究者也报告了各种各样的结果。这说明,有了深度学习,以往计算机所不擅长的图像识别技术得到了很大的发展,从而可以期待计算机能够实现“像人一样”对图像进行处理的技术。

以上两个例子都是将深度学习应用于图像识别的例子。表 1.1 中的最后一个例子是将深度学习应用于语音识别的研究示例。在这一研究中,通过深度学习技术构建语音识别系统,将电话语音识别出来并转换为文字(图 1.3)。

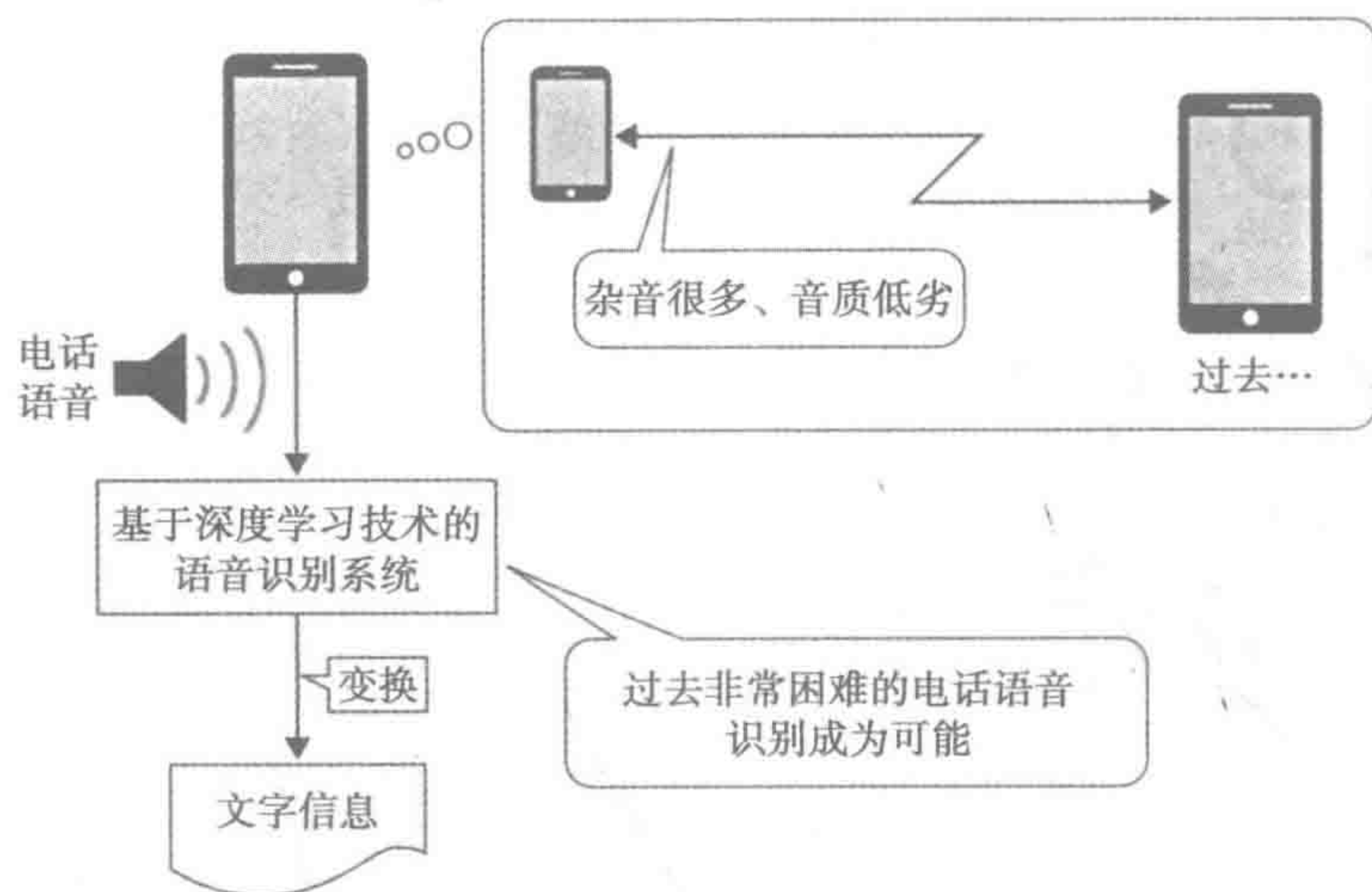


图 1.3 深度学习应用于语音识别系统

语音识别和图像识别一样，是很久之前就开始进行的机器学习研究课题。近年来，在噪声很少的良好环境中进行语音识别，已经达到实用化阶段。然而，当处理的是像电话语音这样的杂音很多、音质很差的对象时，将其内容识别出来并转化为文字就变得非常困难。在这样的条件下，该研究利用深度学习方法构建语音识别系统，获得了传统方法所无法达到的识别精度。该研究不仅仅对语音识别技术的发展做出了贡献，而且说明了深度学习可以作为一种通用的学习方法适用于各种各样的领域。

1.1.2 学习、机器学习和深度学习

深度学习对于机器学习中各个不同领域的研究都产生了非常大的影响。这里，先对学习和机器学习进行整体概述，然后介绍深度学习在其中的地位。

首先，到底什么是学习呢？日常生活中也经常会说到学习。在学校学习是一种典型的学习形态，通过体育或音乐等的练习以习得技能也是关于学习的例子，且在没有明确地指明要学习时也可能发生学习。对工具的使用变得熟练、手套和手的磨合过程等都是隐式的学习成果（图 1.4）。日常生活中不常进行的动作，人经过不断重复也会变得更为擅长，这些都是学习的结果。



图 1.4 各种不同的学习

学习并不是人类所固有的行为，动物也会学习。和人类的学习一样，关于动物的学习也在心理学的各个不同领域有多种多样的研究。

在不管上述哪种场合，通过学习积蓄知识、习得新技能或者使得学习者的经验更为丰富等，都使学习者的内部状态发生了变化。一般来说，这一变化是朝着学习者适应外部环境的方向发展。其结果是，当要处理其他新问题时，学习者能够很巧妙地应付该问题。产生这种状态变化的过程，一般可称为学习。

机器学习是指由机器或者说计算机程序来进行学习。可以说，机器学习和生物的学习本质上是一样的。也就是说，计算机程序通过和外界相互作用，根据其结果变更其内部状态的过程，称为机器学习（图 1.5）。

按照这样的定义，机器学习是范围非常广泛的概念。非常简单的机器学习实例之一是具有变换候补表示功能的日语输入系统，这一学习系统是将英文文字或用片假名输入的日语文字序列变换为片假名和汉字混杂表示的系统。在这一过程中，当输入文字序列对应于多个汉字变换候补时，变换系统会将候补作为序列添加而进行表示。学习系统的

作用是，将过去执行过的变换过程记忆起来，在候补序列中将该变换放到序列的前列来进行表示（图 1.6）。

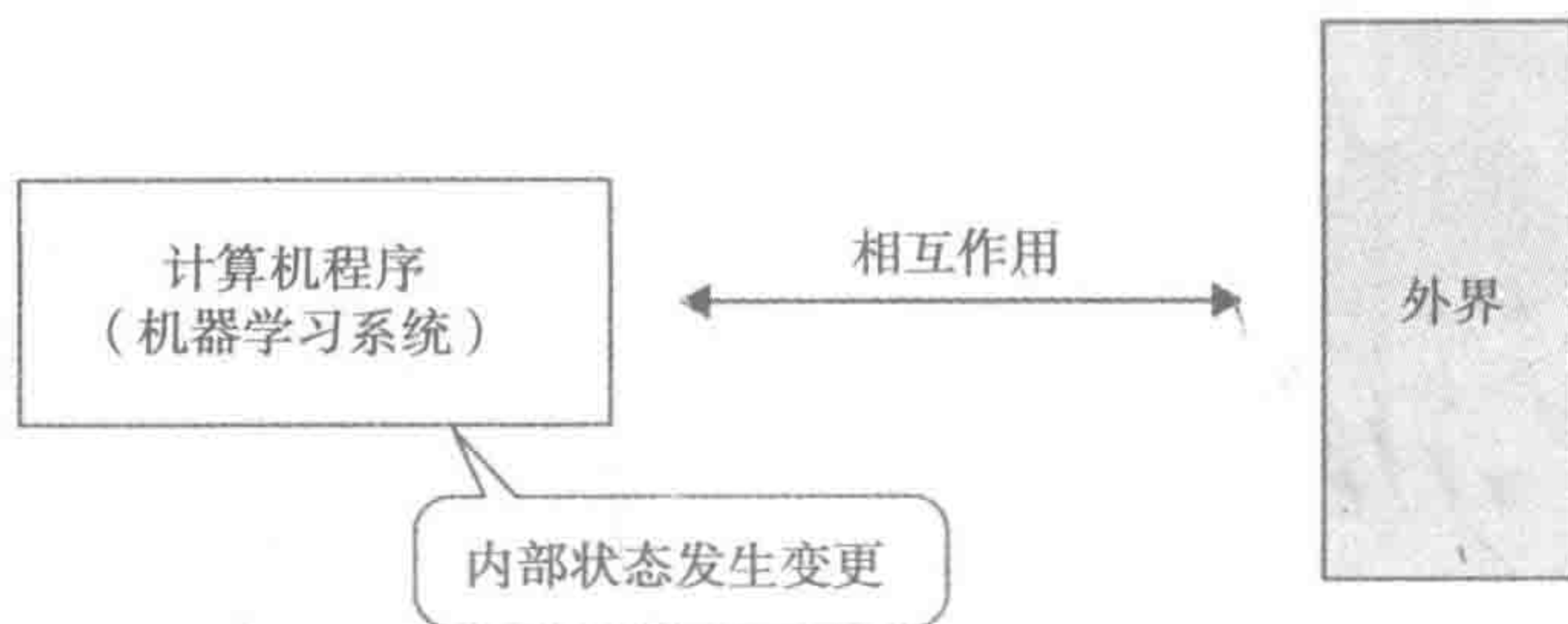


图 1.5 机器学习

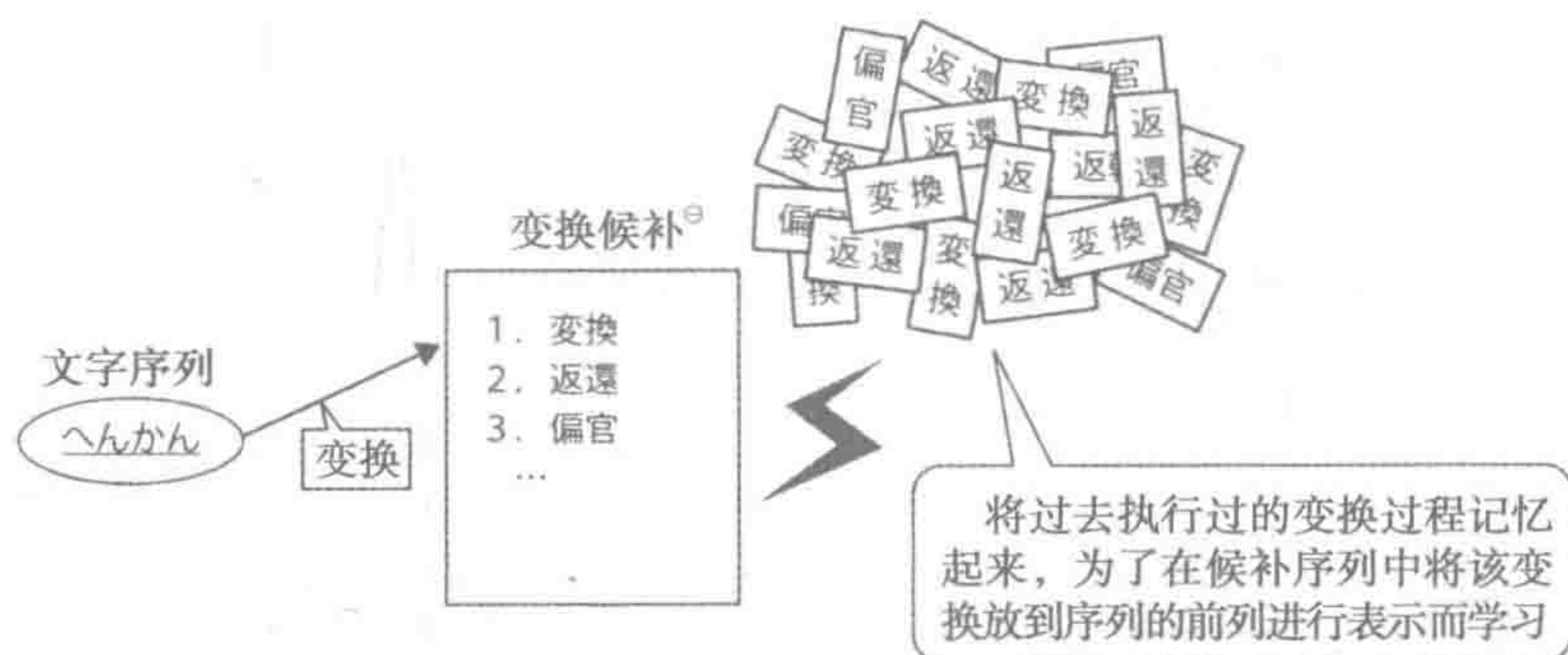


图 1.6 片假名 - 汉字变换系统的机器学习

这一机器学习系统只是将过去的变换结果记忆起来而不做任何其他操作，这是个很方便的系统。这一动作是该系统和“人类”这一外界进行相互作用，根据其结果来生成变换候补的序列，从而变更了系统的内部状态。也就是说，该系统能很好地通过和外界相互作用来获取信息。这就意味着，该系统是机器学习系统。

这种层次的机器学习系统已经组装到了多种装置中。然而，作为学习而言非常重要的泛化（generalization）能力的欠缺，使得像片假名 - 汉字变换系统的候补序列学习这样的例子，只能说是非常简单的学习。

⊖ 方框中的词，都是输入“へんかん”时的日语汉字词。——译者注

泛化是指学习获得的知识和经验的一般化。由于泛化，目前为止所学习的经验能够很合适地对应于不同的新状况。

人们在学校的学习过程中，将学校中学习到的知识作为基础，导出和所学知识有稍许差异的问题或状况的答案。以数学为例，在学习过程中除了有限的练习问题之外没有机会解决其他问题，此时需要将学校的学习经验进行泛化，从而解决初次遇到的其他问题。母语的文章阅读理解也是一样，将学习结果进行泛化，人们能够很肯定地把握第一次读到的文章的内容。这样的“闻一知十”的泛化能力，使得学习的效率和价值非常高。

前面所述的深度学习的例子中，实现了学习的泛化。例如，对于 DQN，对到某个时间节点为止的学习结果进行泛化，就能够应对未知的游戏局面。通过 CNN 进行图像识别的例子中，将和样例不同的测试数据中的图片输入到系统时，凭借学习结果的泛化，能够对该图片进行分类。语音识别的例子也是一样。对于多数机器学习系统而言，它们不仅仅是单纯地记忆过去的事例，而是基于泛化来处理未知状况（图 1.7）。

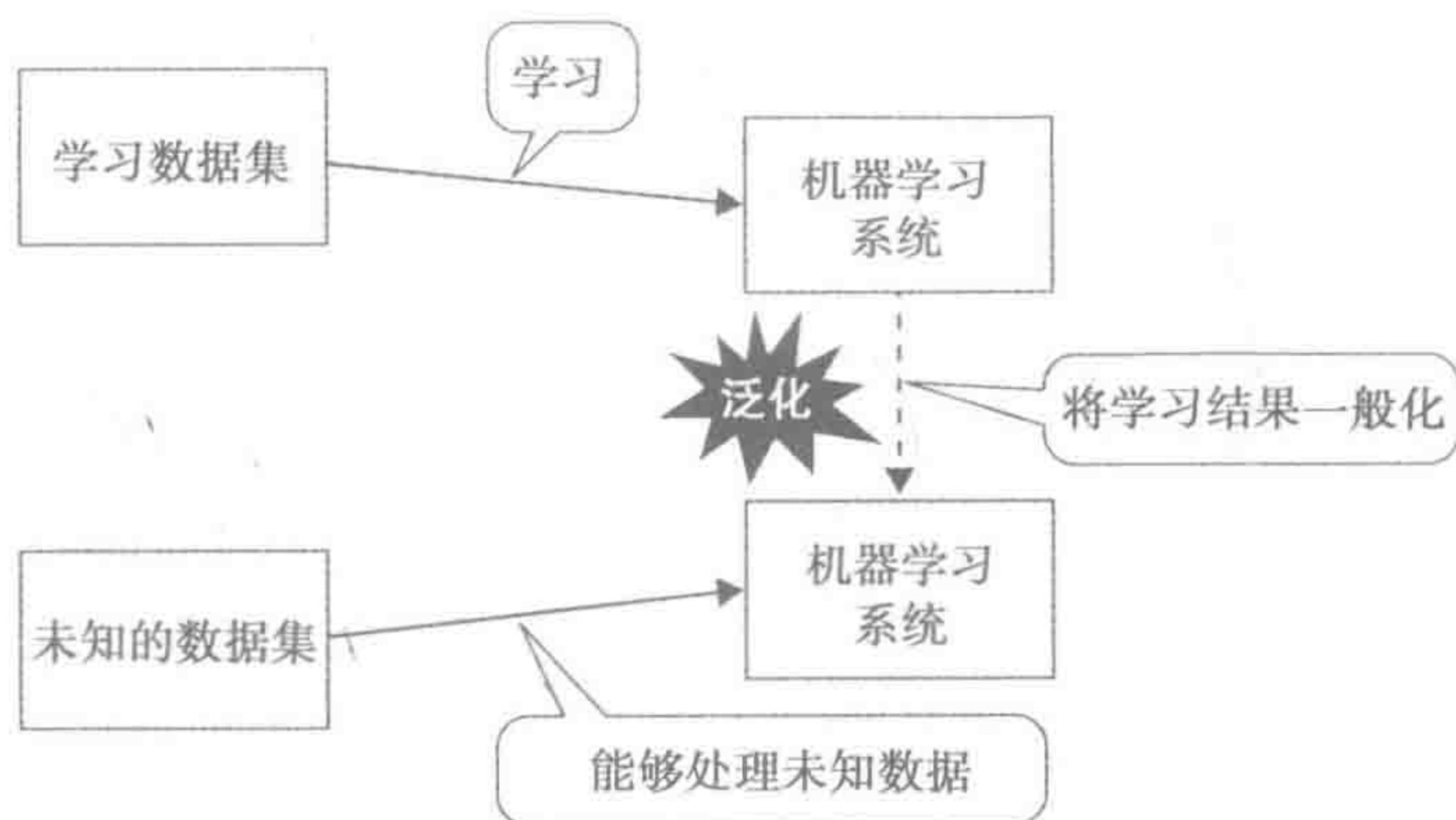


图 1.7 学习和泛化

1.1.3 机器学习的分类

事实上，机器学习的方法各色各样。这里介绍多个机器学习的代表性方法，并介绍深度学习在其中的位置。

机器学习能根据不同的观点进行分类。一种观点认为，学习是基于符号处理的或者是基于统计处理的，可以根据这种观点将机器学习分类（表 1.2）。

表 1.2 机器学习方法的分类（基于处理的原理来分类）

分 类	说 明	例 子
符号处理	以符号处理、符号操作为基础的学习方法	归纳学习 教导学习 演化计算
统计处理	假设学习数据是概率性的数据，以此为主实施数学处理，从而进行学习	统计方法（回归分析、聚类分析、主成分分析等） 神经网络 深度学习

基于符号处理的机器学习的一个很好的例子是，近年来也受到很大关注的大数据（big data）中的文本挖掘（text mining）。大数据是指一般 PC 的硬盘无法存储的海量数据，主要是在互联网上一天天累积起来的数据。而文本挖掘是指，用机器学习方法对大规模文本数据进行处理，从而提取出文本中所含的知识（图 1.8）。

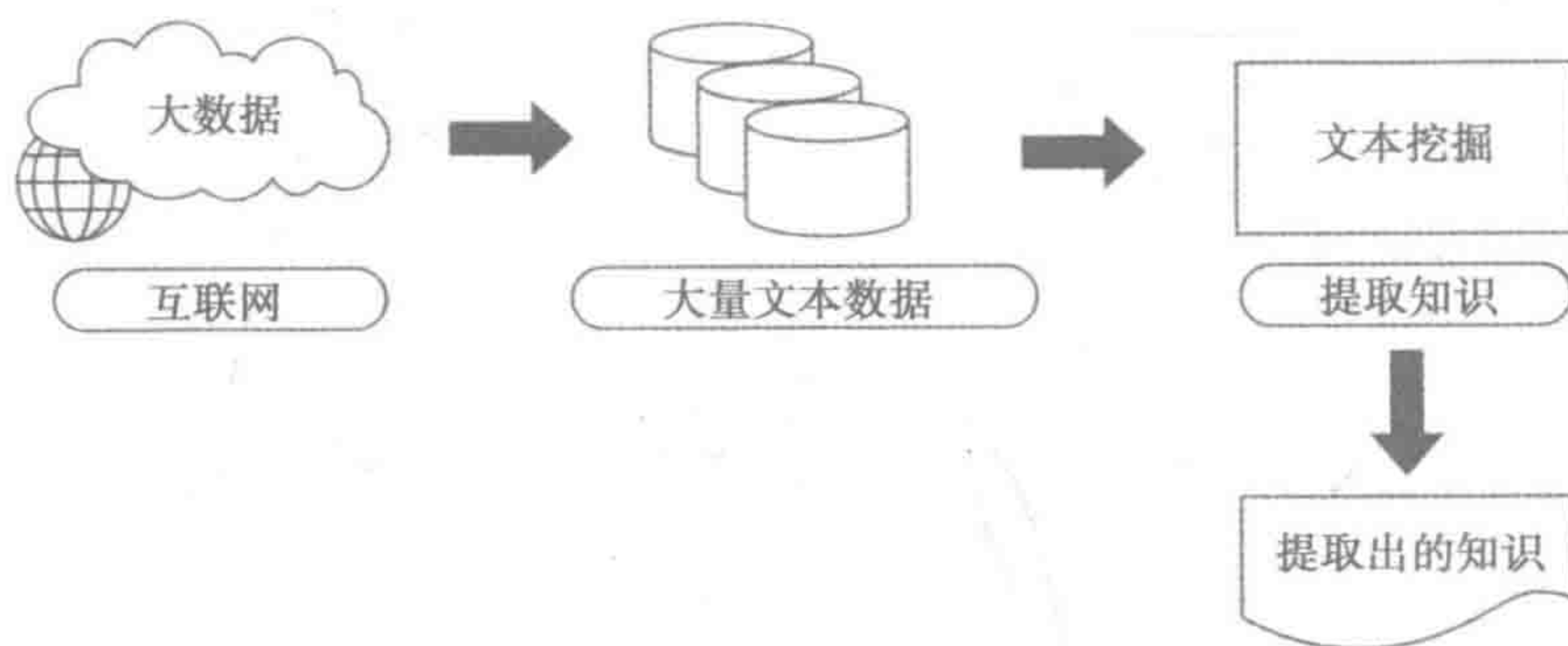


图 1.8 基于文本挖掘的知识获取

在这里，对文本数据进行符号处理，对其结果再次进行符号操作来做分类分析，从而实现机器学习。这些处理中用到了文本处理、自然语言处理、推理、知识表示等基于符号处理的人工智能技术（基于这些技术进行符号处理的机器学习，将在第 2 章中进一步说明）。

称为演化计算 (evolutionary computation) 的机器学习方法是以符号操作为中心进行处理的方法。演化计算是受生物进化思想启发的机器学习方法, 是将生物所具有的遗传结构通过符号处理来实现的学习方法 (关于演化计算, 会在第 3 章中进一步说明)。

对于基于统计处理的机器学习方法, 假设输入的是伴随着误差和噪声的概率性数据, 以此为主实施数学处理进行学习。基于统计学的估计是这方面的典型示例。

此外, 将生物的神细胞模型化的人工神经网络 (Artificial Neural Network, ANN), 也是基于统计处理的机器学习的一个例子。人工神经网络在这一领域中往往用神经网络 (neural network) 来称呼。

将生物神经细胞模型化, 形成人工神经元 (artificial neuron, 称为神经元或神经细胞), 这些人工神经元相互连接形成的网络就是神经网络。人工神经元接收多个输入, 对这些输入进行一定的处理后输出处理结果。这样的处理是非常简单的。基本处理是, 将各个输入值与所对应的系数相乘, 并将所有乘积值相加, 然后将计算结果赋给适当的函数, 这个函数的计算结果就是神经元的输出。这一机制是受生物的神细胞机制的启发而得到的 (图 1.9a)。

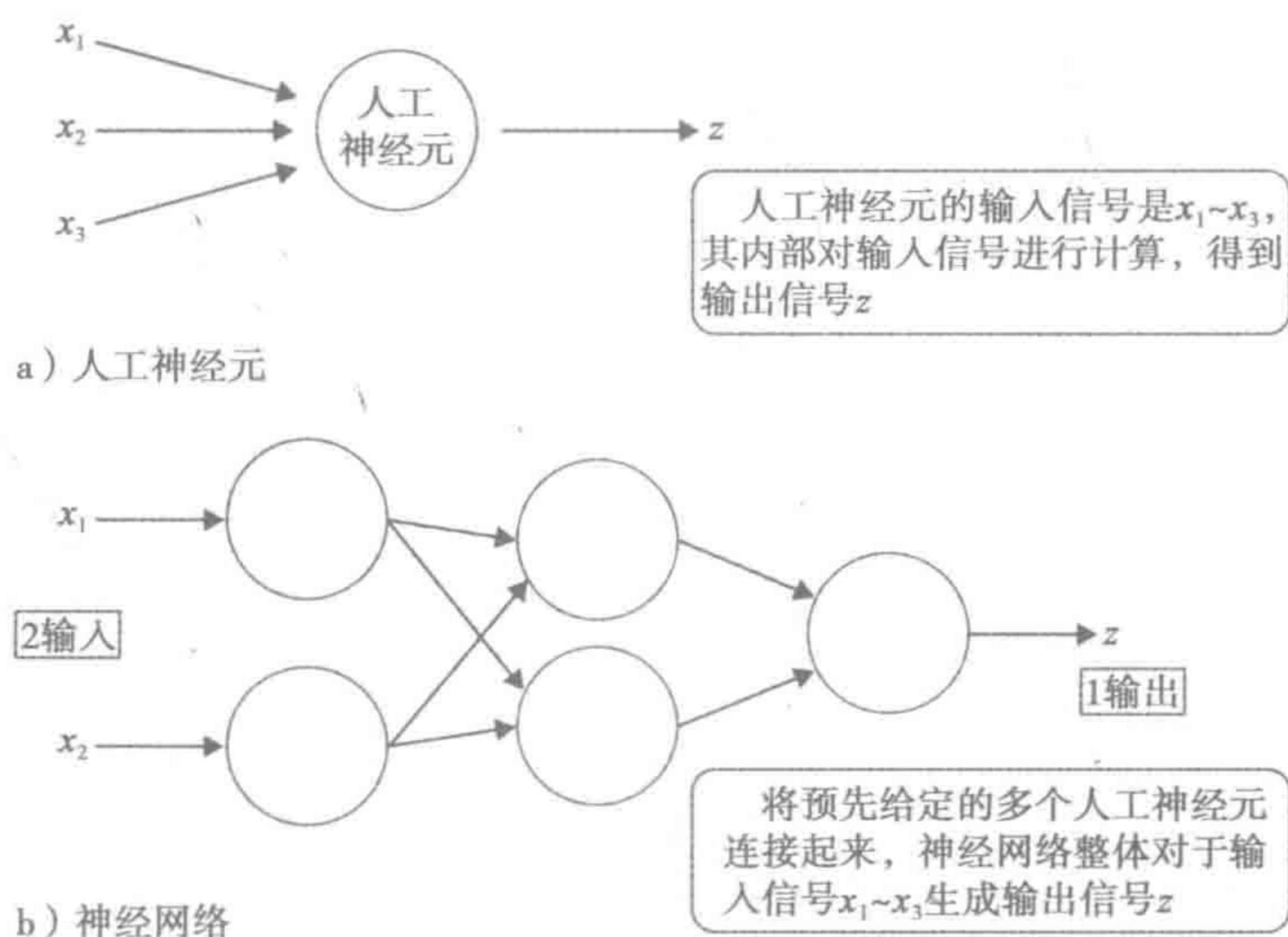


图 1.9 神经网络 (基于统计处理的机器学习的例子)