“十四五”职业教育国家规划教材

工业和信息化精品系列教材

人邮教育

人工智能导论

余明辉 詹增荣 汤双霞◎主编　周珩 谢海燕 时东晓 李莹妮◎副主编
黑马程序员◎组编



AIGC版

人工智能通识课程教材

本书配套丰富的教学资源，
包括教学 PPT、教学大纲、教学设计、
课后习题及答案等

本书详细讲解使用 AIGC 如何实现文本生成、
图像生成、音频生成、视频生成和文档处理，
以及如何进行智能体的开发

中国工信出版集团　人民邮电出版社 POSTS & TELECOM PRESS

目录

单元 1　人工智能与 AIGC　1

1.1　人工智能概述　2
1.1.1　人工智能的基础知识　2
1.1.2　人工智能的核心要素　4
1.1.3　人工智能的主要分支　8
1.2　AIGC 概述　15
1.3　AIGC 的技术实现　17
1.4　常见的 AIGC 创作平台　23
1.5　AIGC 的提示词　27
1.6　人工智能幻觉　30
单元小结　32
课后习题　32

单元 2　AIGC 实践之文本生成　34

2.1　AIGC 的文本生成原理　34
2.2　AIGC 的文本生成工具　37
2.3　AIGC 的文本生成提示词设计　45
实践任务 2-1：科普文章——河流是如何形成的　47
实践任务 2-2：课题报告——中学生课后辅导模式优化　50
单元小结　55
课后习题　55

单元 3　AIGC 实践之图像生成　57

3.1　AIGC 的图像生成原理　57
3.2　AIGC 的图像生成工具　60
3.3　AIGC 的图像生成提示词设计　67
实践任务 3-1：端午节海报　77
实践任务 3-2：电商宣传图　80
单元小结　83
课后习题　83

单元 4　AIGC 实践之音频生成　85

4.1　AIGC 的音频生成原理　85
4.2　AIGC 的音频生成工具　87
4.3　AIGC 的音频生成提示词设计　90
实践任务 4-1：景点讲解语音　95
实践任务 4-2：青年之歌　97
单元小结　98
课后习题　98

单元 5　AIGC 实践之视频生成　100

5.1　AIGC 的视频生成原理　100
5.2　AIGC 的视频生成工具　102
5.3　AIGC 的视频生成提示词设计　106
实践任务 5-1：立冬视频　110
实践任务 5-2：新车发布宣传视频　112
单元小结　113
课后习题　114

单元 6　AIGC 实践之文档处理　116

6.1　AIGC 的文档处理原理　116
6.2　AIGC 的文档处理工具　119
6.3　文档处理的要点　123

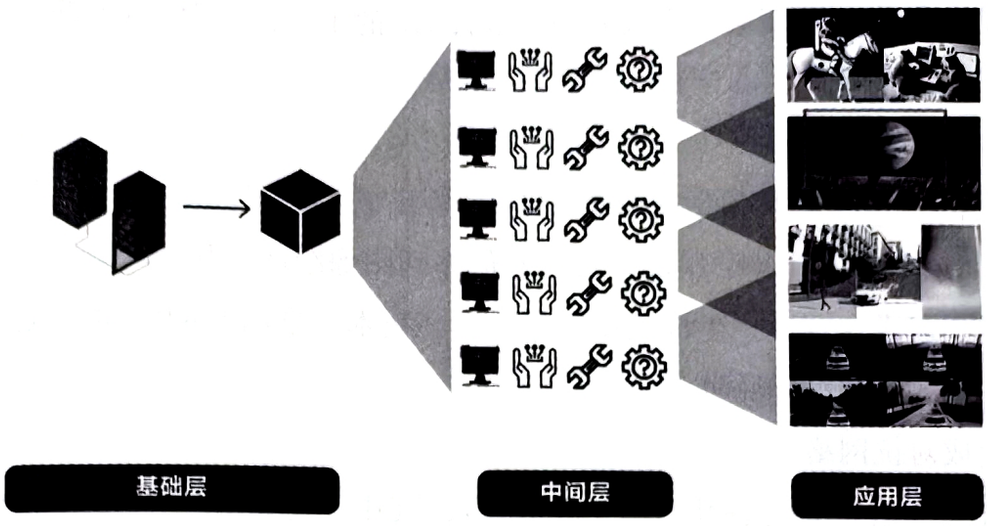


图 1-4　AIGC 的产业生态体系

在AIGC的产业生态体系中，基础层、中间层和应用层相互协作，共同推动AIGC技术的发展与应用落地，这3层的说明如下。

（1）基础层

基础层是AIGC产业生态体系的基石，为整个体系提供底层算力资源和数据资源。算力资源涵盖AI芯片、存储、网络、安全以及基础设施软件等；数据资源则包括AIGC模型训练和优化所需的大量高质量多模态数据。

（2）中间层

中间层是AIGC产业生态体系的核心，该层主要基于基础层大规模预训练模型，开发出针对特定行业或应用场景的小型化、定制化模型。同时，中间层提供给开发者可使用的各种应用程序接口（Application Program Interface，API）、软件开发工具包（Software Development Kit，SDK）、在线服务平台，以及数据标注、模型优化等服务，便于开发者快速部署和集成AIGC功能，更好地满足不同场景的需求。

（3）应用层

应用层是AIGC产业生态体系的终端，涵盖针对企业的专用模型应用和针对个人用户的个性化服务。在企业领域，AIGC可应用于传媒、电商、影视等多个行业；在个人用户领域，AIGC则提供了图像创作、音乐制作、视频剪辑等个性化服务，让个人用户能够轻松享受到AIGC带来的便利和乐趣。

1.3　AIGC 的技术实现

AIGC的应用领域较为广泛，其核心在于利用生成式模型的理论和技术来实现各类

内容的自动化创作。下面基于生成式模型和AIGC的工作流程对AIGC的技术实现进行说明。

1. 生成式模型

生成式模型是能够学习数据分布并依据学习成果生成新数据的模型。近年来，AIGC技术如风暴般席卷各个领域，从文本创作到图像绘制，从音频合成到视频生成，它无处不在。而在这背后，生成式模型作为核心技术，推动着AIGC的快速发展。下面对常见的生成式模型进行说明。

（1）生成对抗网络

生成对抗网络主要包括生成器和判别器，其中，生成器通常由多层神经网络组成，它接收一个随机噪声向量作为输入，通过一系列的变换将随机噪声向量映射为数据样本，如图像、音频等。判别器通常也由多层神经网络组成，它的输入是生成器生成的样本和真实样本，输出是一个表示样本真实性的概率值。在训练过程中，生成器的目标是生成能够欺骗判别器的样本，而判别器的目标是准确区分真实样本和生成样本。通过不断的对抗训练，生成器和判别器的性能都能得到提升，最终生成器能够生成高质量的样本。生成对抗网络的训练过程如图1-5所示。

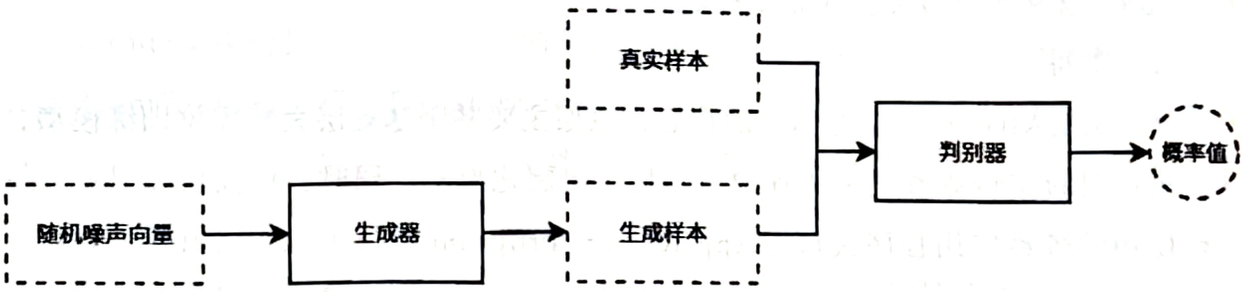


图 1-5　生成对抗网络的训练过程

生成器根据传入的随机噪声向量生成样本；判别器接收真实样本和生成样本进行判别，判别后会生成一个介于0和1之间的数字，代表样本为真实样本的概率。1表示判别器确定样本是实样本，0表示判别器判定样本是由生成器伪造的。生成对抗网络的训练过程是一个不断迭代的过程，生成器和判别器会根据判别的结果不断自我改进。生成器不断调整参数，力求使生成的样本更接近真实样本，使判别器难以判别；判别器则持续优化自身，提升对真假样本的分辨能力。通过反复的对抗训练，二者的性能持续提升，直至生成器生成的样本与真实样本极为相似，此时达到理想的生成效果。

当前生成对抗网络在图像生成领域取得了巨大的成功，它可以生成逼真的人脸图像、风景图像、艺术画作等。此外，生成对抗网络还被应用于视频生成、语音合成、虚拟场景构建等。

（2）变分自编码器

变分自编码器是机器学习中的一种生成模型，它能够基于训练数据生成新的数据变体，其核心思想是将数据编码到一个潜在空间中，然后从这个潜在空间中解码出数据，使得解码后的数据能够尽可能地还原为原始输入数据。与传统自编码器不同的是，变分自编码器对潜在空间的分布进行了约束，使其服从先验分布（通常是高斯分布）。这样做的目的是让潜在空间具有更强的连续性和泛化能力，从而使模型能够生成更加多样化和合理的数据样本。变分自编码器的训练过程如图1-6所示。

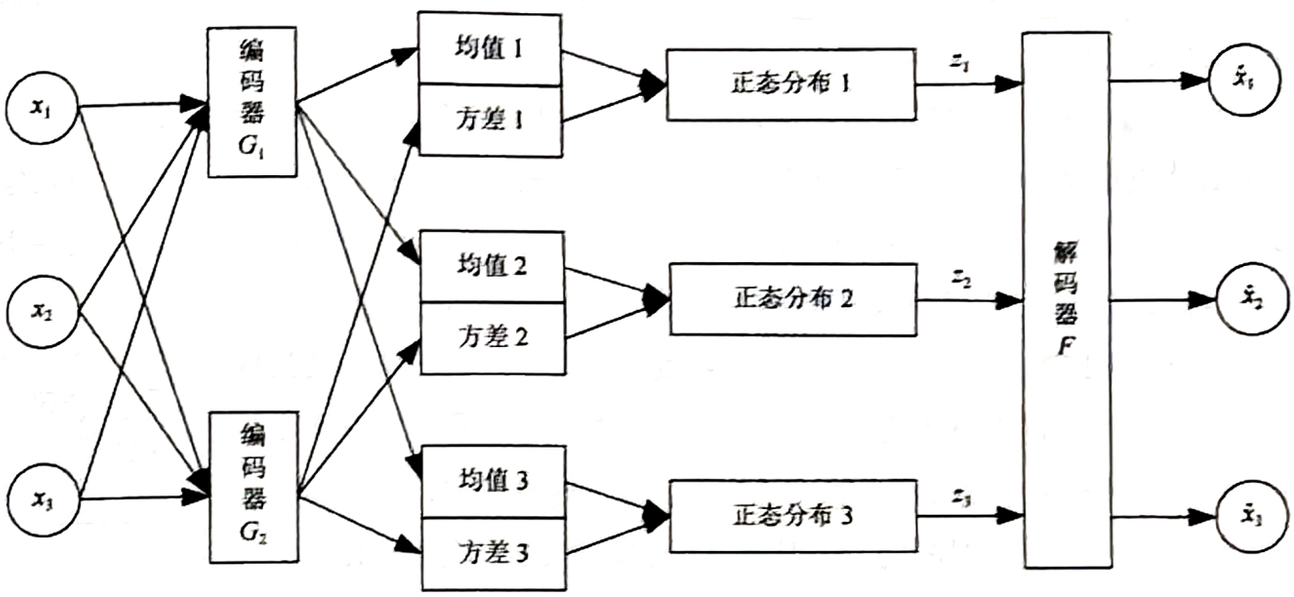


图 1-6　变分自编码器的训练过程

首先，编码器接收到输入数据后会将其映射到潜在空间，从而得到潜在变量概率分布的参数：均值和方差。接着基于得到的均值和方差，从标准正态分布里采样，得到潜在空间样本。最后，解码器接收潜在空间样本，并且对这个潜在空间样本进行一系列的转换和处理，以生成与原始输入数据高度相似的数据，让重构后的数据在特征和信息上尽可能逼近原始输入数据。

变分自编码器的训练过程中，模型会计算两个关键的损失：一个是重构损失，它衡量的是解码器重构出的数据与原始输入数据之间的差异程度；另一个是KL散度（Kullback-Leibler Divergence，它是由库尔贝克和莱布勒提出的一个用于衡量两个概率分布之间差异的非对称度量）损失，用于约束潜在变量的分布，使其尽量接近标准正态分布。将这两个损失组合起来得到总损失，模型会利用反向传播算法自动计算出每个参数应该如何调整才能让总损失最小，之后会使用诸如随机梯度下降之类的优化算法对编码器和解码器的参数进行更新。不断重复上述步骤，持续迭代优化模型，直到模型具有较好的性能，能够稳定地生成高质量的重构数据。

当前变分自编码器在图像生成、图像编辑、数据压缩等方面都有广泛的应用。例如，通过在潜在空间中进行操作，变分自编码器可以实现图像的风格迁移、属性编辑等功能。

（3）扩散模型

扩散模型的名字来源于自然界的扩散现象，扩散模型通过逆转扩散过程来生成新数据。简单来说，扩散模型通过逐步向数据中添加按特定概率分布生成的随机噪声，使数据逐渐变为纯噪声，再逆转这一过程，使纯噪声变为与原始数据分布相似的新数据。从上述内容可以得出，扩散模型主要有正向扩散和逆向扩散两个过程，下面对这两个过程进行说明。

① 正向扩散。扩散模型的正向扩散是一个逐步将数据转换为噪声的过程。首先，系统会从数据集中随机选取一份样本数据，例如一张猫的图片；接着，系统按照预定的噪声调度策略，逐步向这份样本数据中添加噪声。每一次添加噪声后都会生成一个新版本的数据，逐渐破坏原始数据的信息，直到最终得到完全由噪声组成的数据。

整个正向扩散过程根据预设的噪声调度策略系统性地将清晰的数据转换为纯噪声，这个过程为后续的逆向扩散过程提供了明确的学习目标，即如何从噪声中恢复出有意义的数据。通过这种方式，模型能够具有去除噪声、重建数据的能力，从而降低实际生成（逆向扩散）的难度，并提高了生成结果的质量。正向扩散的过程如图1-7所示。



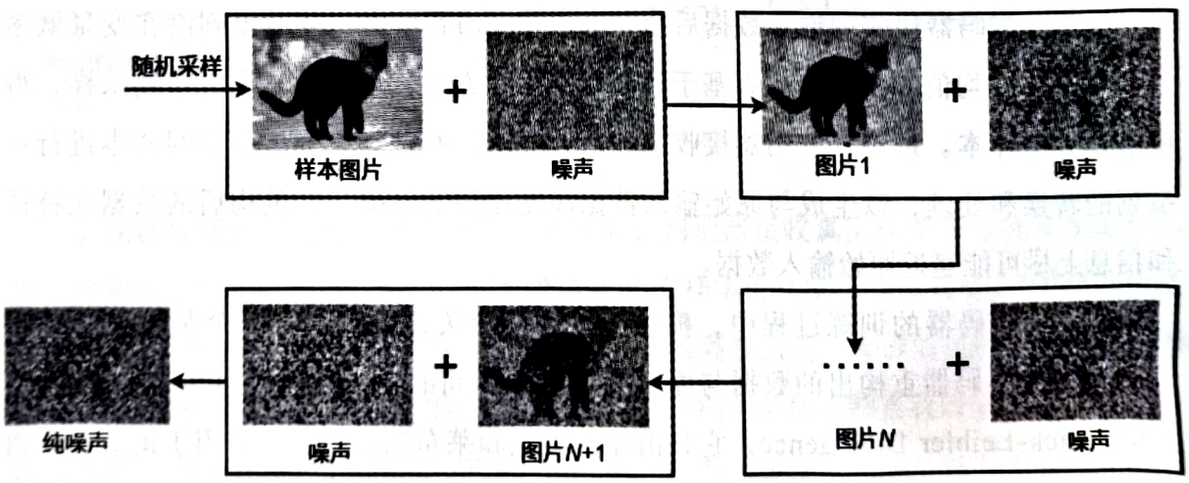


图 1-7 正向扩散的过程

② 逆向扩散。扩散模型的逆向扩散是一个从噪声中恢复数据的过程，系统会将纯噪声作为输入逐步迭代去噪。每一步都会利用训练好的神经网络预测当前噪声中的原始信息残留，通过计算生成更接近真实数据分布的新版本数据，逐步重建细节（如物体轮廓、色彩纹理）。

整个逆向扩散过程依托在正向扩散中学习到的噪声分布规律，通过迭代消除噪声干扰，逐步细化局部特征。这一过程将复杂的图像生成任务拆解为连续的噪声消除子任务，模型通过学习正向扩散中“数据如何演变为噪声”的规律，反向推导出“噪声如何恢复为数据”的路径。最终生成的数据不仅符合真实数据分布，还能根据语义输入引入创造性变化，实现从无到有的内容生成。逆向扩散的过程如图1-8所示。

【扫码看图】

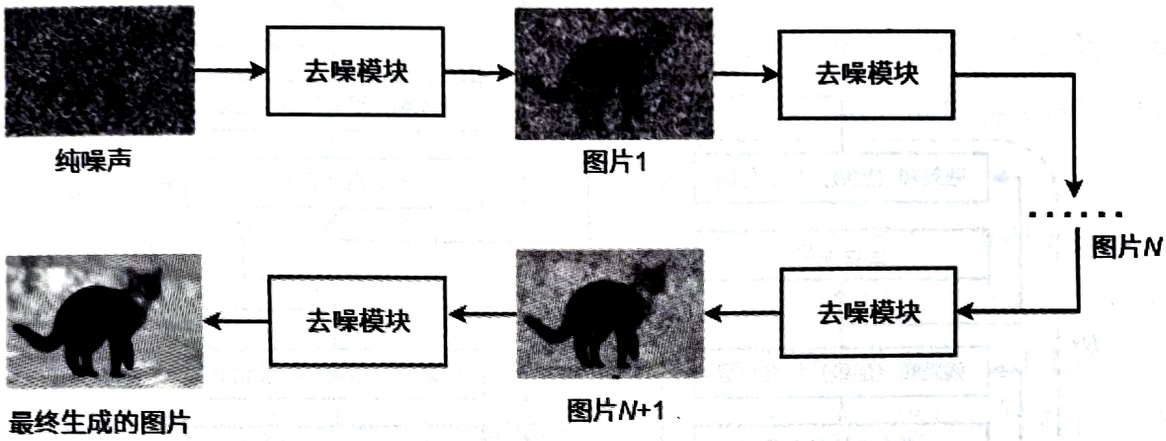


图 1-8　逆向扩散的过程

（4）基于Transformer的模型

Transformer是一种基于注意力机制的深度学习模型架构，它被设计用于处理序列数据，特别是在自然语言处理领域取得了巨大的成功。与传统的循环神经网络和长短期记忆网络相比，Transformer能够更有效地捕捉序列中的长距离依赖关系，并且由于其具有并行处理能力，训练速度也更快。

Transformer通过计算输入序列中每个位置的注意力权重来动态地关注不同位置的信息。基于Transformer的预训练语言模型在大规模语料上进行无监督预训练，学习语言的语法、语义和语用知识。在预训练之后，可以通过微调将模型应用于各种自然语言处理任务，如文本生成、问答系统、机器翻译等。从结构上看，Transformer主要由编码器和解码器组成，Transformer的结构如图1-9所示。

图1-9中，左边虚线框住的内容表示编码器，*Nx*表示有*N*个编码器；右边虚线框住的内容表示解码器，*Nx*表示有*N*个解码器。其中，编码器负责将输入序列编码成一个连续的向量表示，解码器则根据编码器的输出以及上轮输出词编码后的信息生成下一个输出。

当前基于Transformer的模型在文本生成领域表现出色，能够生成高质量的文章、对话等。同时，该类模型在智能客服、知识图谱构建、信息检索等领域也有广泛的应用。

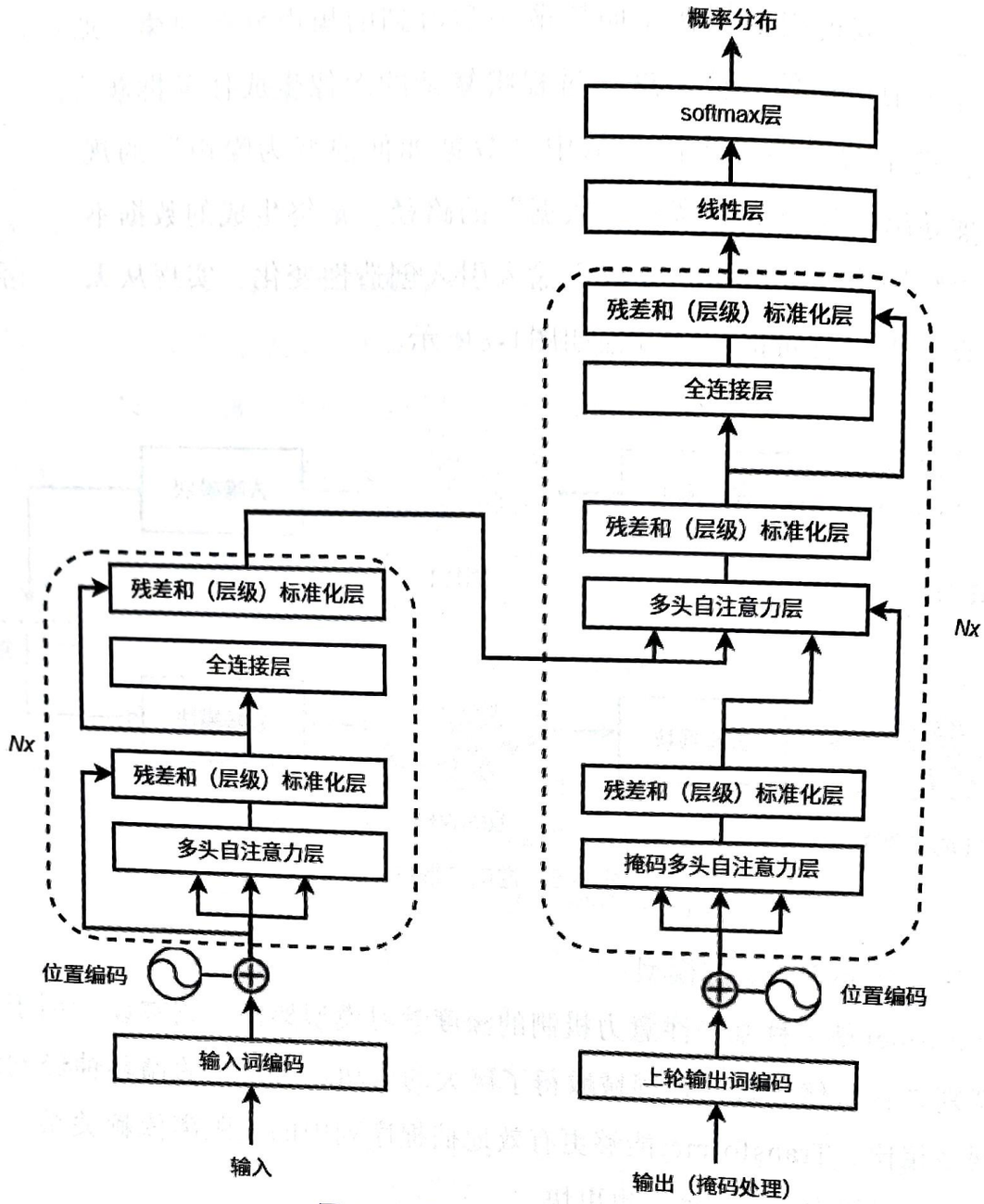


图 1-9 Transformer 的结构

书名	书号
Java 基础案例教程（第 3 版）	978-7-115-65384-0
Java Web 程序设计任务教程（第 3 版）（AIGC 版）	978-7-115-66552-2
Java EE 企业级应用开发教程（Spring+Spring MVC+MyBatis）（第 3 版）	978-7-115-66565-2
Java EE 企业级应用开发项目教程（SSM）	978-7-115-60347-0
微服务架构基础（Spring Boot+Spring Cloud+Docker）（第 2 版）	978-7-115-66572-0
Spring Boot 企业级开发教程（第 2 版）	978-7-115-63438-2
Spring Cloud 微服务架构开发（第 2 版）	978-7-115-64609-5
C# 程序设计基础入门教程（第 3 版）	978-7-115-68197-3
C 语言程序设计案例式教程（第 2 版）	978-7-115-59863-9
Python 快速编程入门（第 3 版）	978-7-115-64973-7
Python 程序设计现代方法（第 2 版）	978-7-115-63655-3
Python 数据可视化（第 2 版）	978-7-115-65656-8
Python 数据预处理	978-7-115-56206-7
Python 网络爬虫基础教程（第 2 版）	978-7-115-68925-2
Python Web 开发项目教程（Flask 版）	978-7-115-60020-2
软件测试（第 2 版）	978-7-115-61638-8
HTML5+CSS3 网页设计与制作（第 2 版）	978-7-115-65341-3
JavaScript+jQuery 交互式 Web 前端开发（第 2 版）	978-7-115-63379-8
Vue.js 前端开发实战（第 2 版）	978-7-115-61815-3
Bootstrap 响应式 Web 开发（第 2 版）	978-7-115-64572-2
HTML+CSS+JavaScript 网页制作案例教程（AI 助学）（第 3 版）	978-7-115-68926-9
JavaScript 前端开发案例教程（第 2 版）	978-7-115-59323-8
HTML5+CSS3 网站设计基础教程（第 3 版）	978-7-115-61406-3
网页设计与制作项目教程（HTML+CSS+JavaScript）（第 2 版）	978-7-115-56598-3
响应式 Web 开发项目教程（HTML5+CSS3+Bootstrap）（第 3 版）	978-7-115-65706-0
PHP 网站开发实例教程（第 3 版）	978-7-115-63654-6
PHP+MySQL 动态网站开发（第 2 版）	978-7-115-64765-8
Web 前端模块化开发教程（ES6+Node.js+Webpack）	978-7-115-56462-7
微信小程序开发实战（第 2 版）	978-7-115-60602-0
Android 移动开发基础案例教程（第 3 版）	978-7-115-68026-6
自动化测试应用教程（Web+App）	978-7-115-59768-7
Laravel 框架开发实战	978-7-115-56326-2
计算机网络技术及应用（第 2 版）	978-7-115-66563-8
Photoshop CC 设计与应用任务教程（第 3 版）	978-7-115-67277-3
Linux 网络操作系统项目化教程（openEuler）	978-7-115-68310-6
Linux 网络操作系统项目化教程（麒麟版）	978-7-115-68331-1
Linux 网络操作系统项目化教程（Ubuntu）	978-7-115-69118-7
鸿蒙 HarmonyOS 应用开发基础	978-7-115-66566-9

用书教师扫码下载
本书配套资源

人邮教育
www.ryjiaoyu.com
教材服务热线：010-81055256
反馈 / 投稿 / 推荐信箱：315@ptpress.com.cn
人邮教育服务与资源下载社区：www.ryjiaoyu.com

ISBN 978-7-115-68883-5
9 787115 688835 >

定价：49.80 元