



第二章

机器学习基础

2.4 从机器学习到深度学习

课程名称: 人工智能通识教程

主讲人:

日期:

单位: 华南农业大学



目录

CONTENTS

- 01 知识回顾
- 02 案例引入
- 03 传统机器学习与深度学习对比
- 04 常见深度学习模型介绍
- 05 常见深度学习框架介绍
- 06 课堂总结

重点难点



从机器学习到深度学习

回顾：什么是机器学习？

案例引入

传统机器学习的局限性

引入深度学习

传统机器学习与深度学习对比

常见深度学习模型介绍

常见深度学习框架介绍



机器学习（Machine Learning, ML）是一个多学科交叉领域，研究计算机模拟或实现人类的学习行为、自动从数据中学习规律的过程。机器学习是一种从观察到的数据（样本）中发现规律，并利用学到的规律（模型）对未观察到的位置和数据进行预测的方法，是人工智能的一类基础技术。

知识回顾：什么是机器学习



传统的机器学习主要分为以下几类

有监督学习（Supervised Learning）：每个样本都有对应的目标值（标签），模型通过学习输入特征与标签之间的关系来进行预测。常见的任务包括回归（预测连续值）和分类（预测离散类别）。

无监督学习（Unsupervised Learning）：样本没有标签，模型的目标是从数据中发现潜在的结构或模式，常见的任务包括聚类和降维。

半监督学习（Semi-supervised Learning）：介于有监督和无监督学习之间，部分样本有标签，部分样本没有标签。

机器学习三要素：模型（Model）、学习准则（Learning Criterion）、优化算法（Optimization）

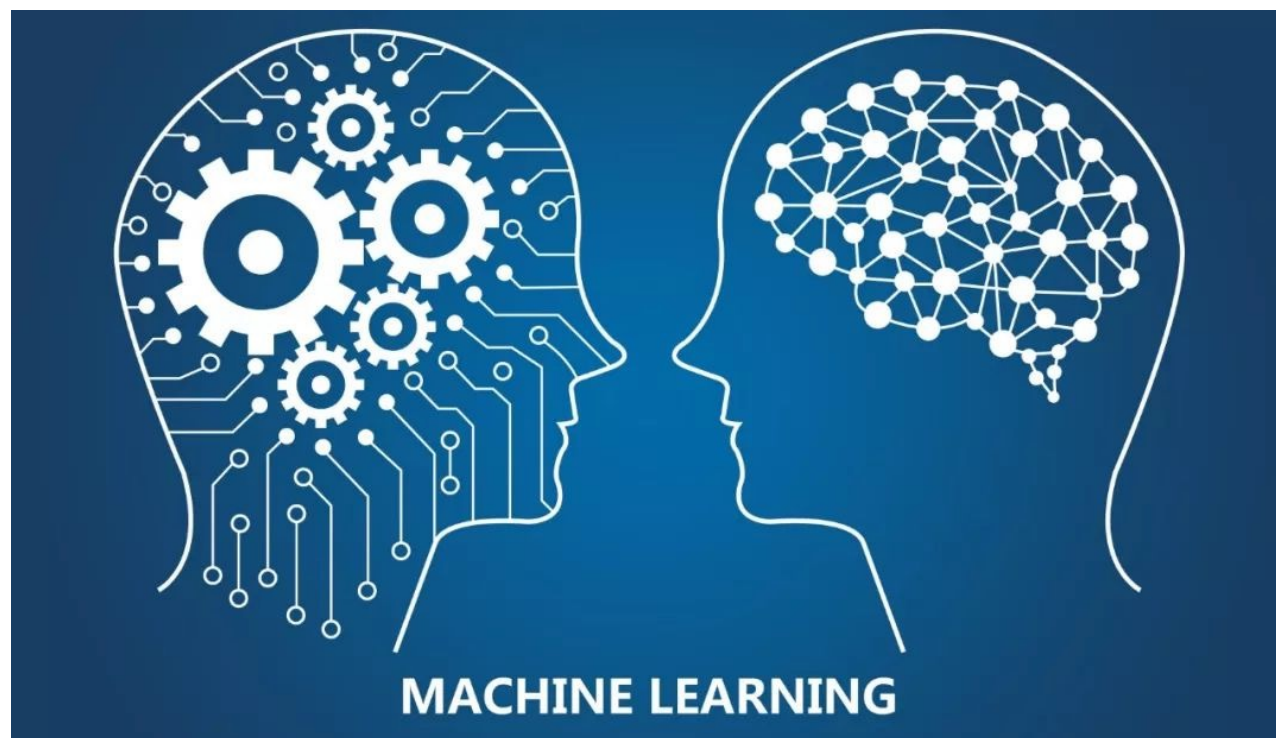
机器学习模型是用于从数据中学习规律并进行预测的工具。模型本质上是一个“函数”，把我们给它的“输入”（数据）转化为想要的“输出”（标签）。根据任务的不同，机器学习模型可以分为几大类，包括分类模型、回归模型和聚类模型等。



学习准则是用来衡量模型预测结果与真实值之间差异的标准。在有监督学习中，学习准则通常通过损失函数（Loss Function）或代价函数（Cost Function）来定义。

机器学习中使用的任何损失函数都会变成寻找最小值的问题（如果损失函数越大越精确，可以先乘以负号）。优化算法是用来高效地找到最优模型参数的方法，即通过调整参数 θ ，使得损失函数的值最小化。其核心任务是找到使损失函数最小的模型参数 θ ，从而让模型的预测结果尽可能接近真实值。

为什么需要从机器学习到深度学习？





案例引入

案例引入

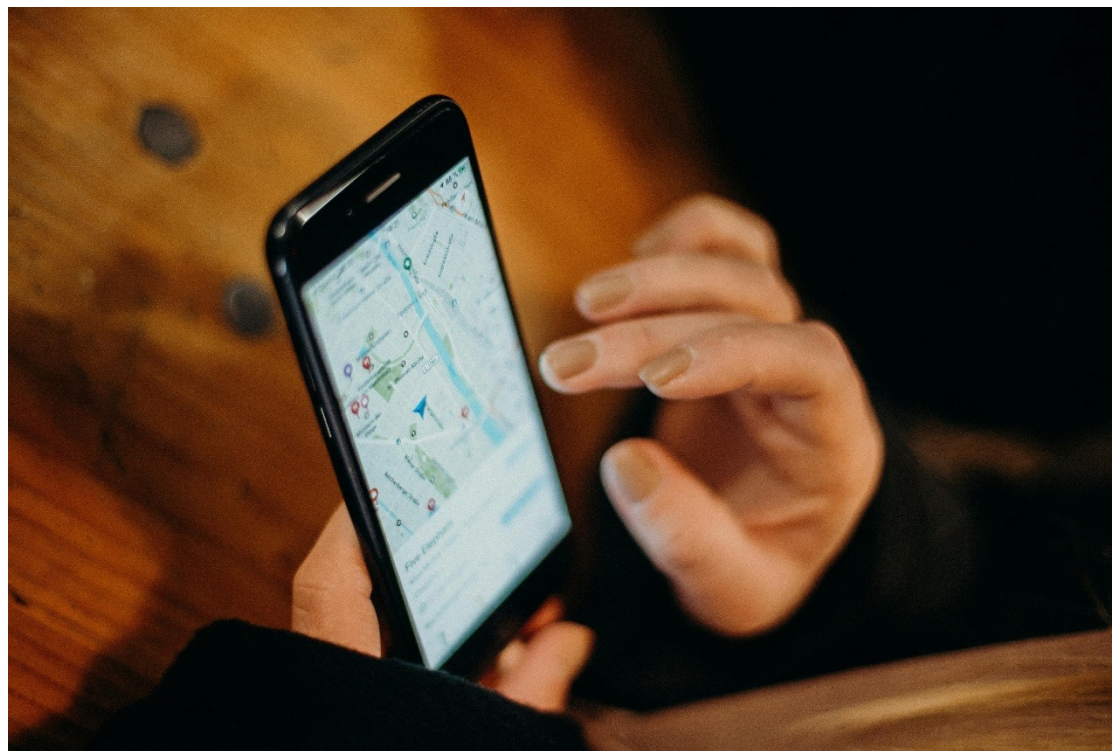


城市交通拥堵是全球大城市面临的共同难题，在我们日常生活中，高峰期堵车几乎成了“家常便饭”。

堵车不仅浪费我们的时间、影响心情，还会增加了空气污染、油耗。



假设我们能提前知道哪条路会堵车，是不是就可以选择一条更顺畅的路线？那么，有没有办法“预测”未来的交通状况呢？



我们可以尝试用传统机器学习来预测堵车。我们来设想一个简单的交通预测系统，目标是：

根据当前的天气、时间、路段等信息，预测某条道路是否会堵车。

传统方法的第一步是：人工设计哪些因素会影响交通状况。所以我们需要收集一些数据，用于传统机器学习模型的训练。



思考与讨论

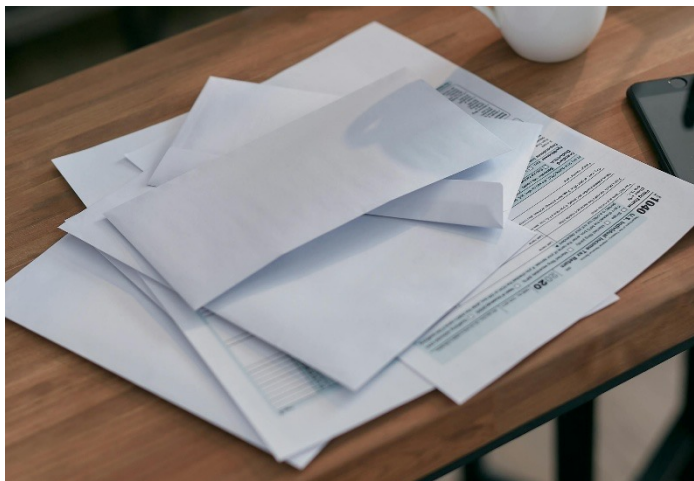
哪些因素会影响交通状况？



影响因素实例

影响因素（输入）	说明
时间段	不同时间段车流量不同
天气情况	晴天、雨天、雪天？恶劣天气容易堵车
是否节假日	节假日车少，平时上下班高峰期车多
是否有施工	路段是否封闭？是否有施工？影响通行
是否有事故	最近有没有发生交通事故？

传统机器学习模型本身不便于从原始数据中提取信息，所以原始数据需要人工的整理设计，让采集的数据变成模型可以方便“读懂”的特征。



原始数据



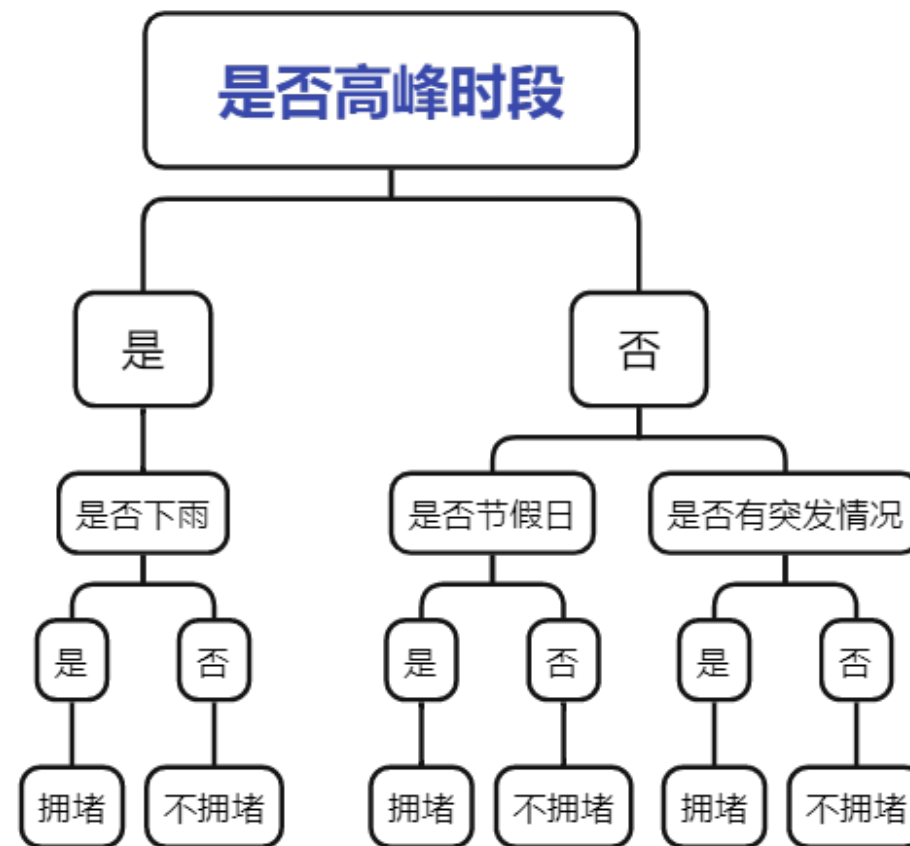
人工数据处理



特征数据

接下来，我们需要用这些特征来训练一个简单的机器学习模型。因为我们的目标是预测道路是否会堵车，所以可以采用传统机器学习中的分类模型来进行训练。比如：决策树、支持向量机（SVM）等模型

这些模型就像是一本“规则手册”，告诉计算机：“如果满足A、B、C条件，就认为会堵车；否则不会。”



决策树模型示意图

案例引入



实际上，交通系统非常复杂，涉及很多因素：

- 天气变化
- 车流量波动
- 交通事故
- 道路施工
- 驾驶员行为……

我们不可能把所有情况都人工写进规则里，而且数据时时刻刻都在产生，数据量太大，人工根本来不及处理，模型也很难及时更新。



正如案例中所见，随着时代的发展，传统机器学习逐渐暴露出一些瓶颈：特征工程依赖人工设计，效率低且难以覆盖复杂场景；同时，模型本身的复杂度受限，难以捕捉数据中的深层规律。

当传统机器学习遇到瓶颈时，我们是否还能找到新的出路？



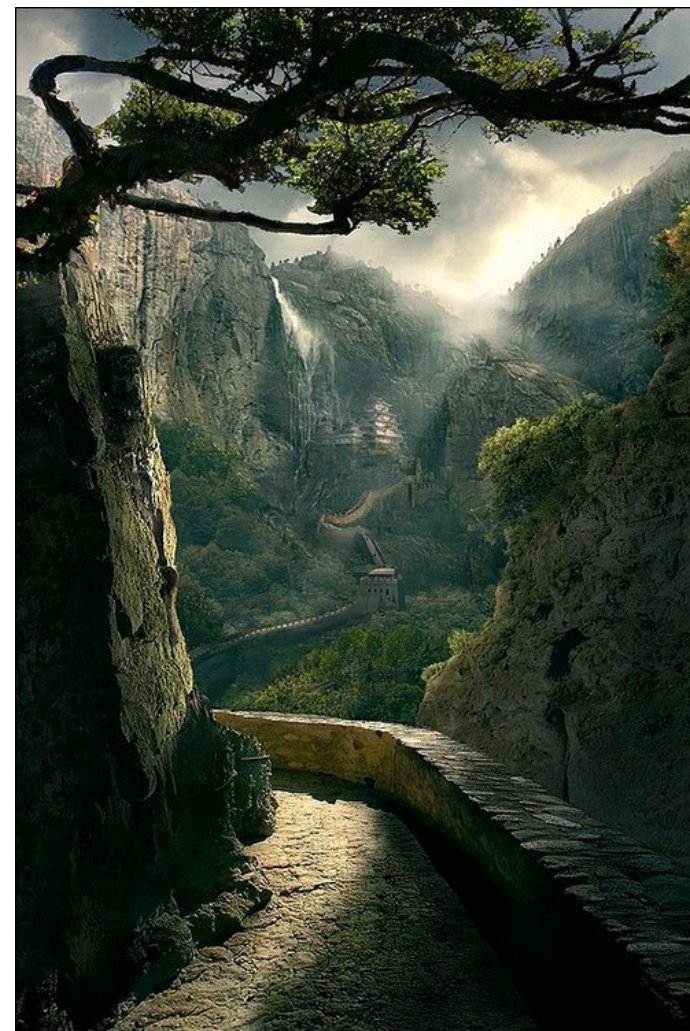
山重水复疑无路

案例引入



科技的持续进步，特别是在计算能力的显著提升、海量数据的积累以及核心算法的创新突破，极大地推动了深度学习的发展。相较于传统方法，深度学习能够实现特征的自动学习与高效表达，显著增强了对高维、复杂数据的建模与处理能力。

深度学习的本质：
使用“神经网络”模拟人脑的学习机制
利用多层结构逐层提取抽象特征



山重水复疑无路，
柳暗花明又一村。
——陆游《游山西村》

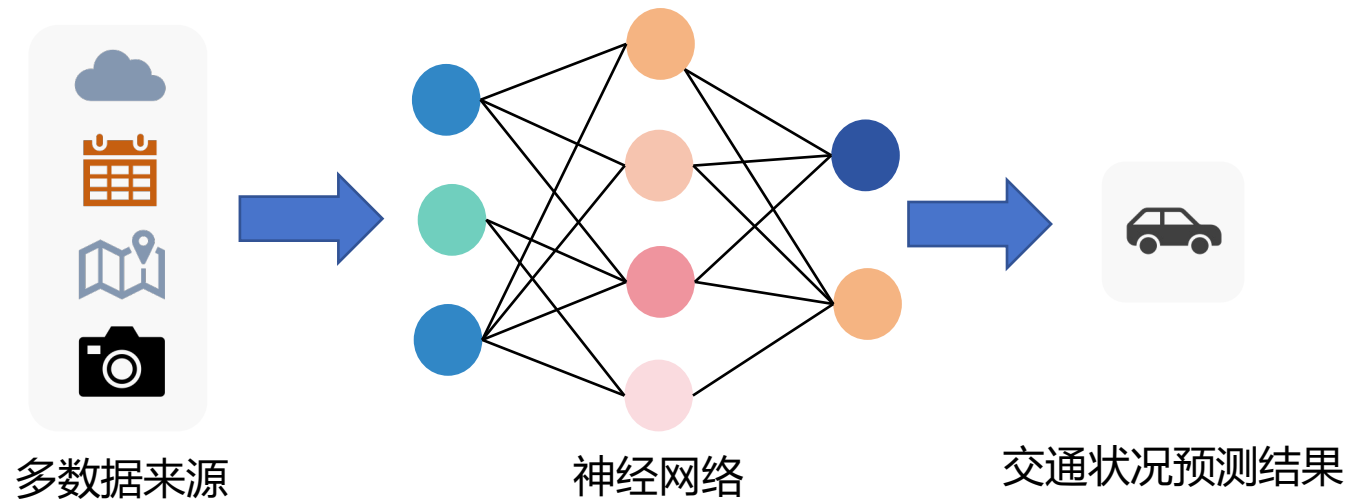


思考与讨论

如何用深度学习来解决交通拥堵预测的问题？

深度学习能够从海量的历史与实时数据中自动挖掘出影响交通流量的关键因素，并建立复杂的非线性关系模型。相比传统方法依赖人工设定规则的方式，深度学习具备更强的自适应性和泛化能力。

深度学习采用“端到端”的建模方式，即直接将原始输入（如传感器采集的数据）映射到输出结果（如未来某时段的拥堵概率），无需人工设计特征或制定规则。这种自动化流程不仅提升了效率，也减少了人为偏见的影响。



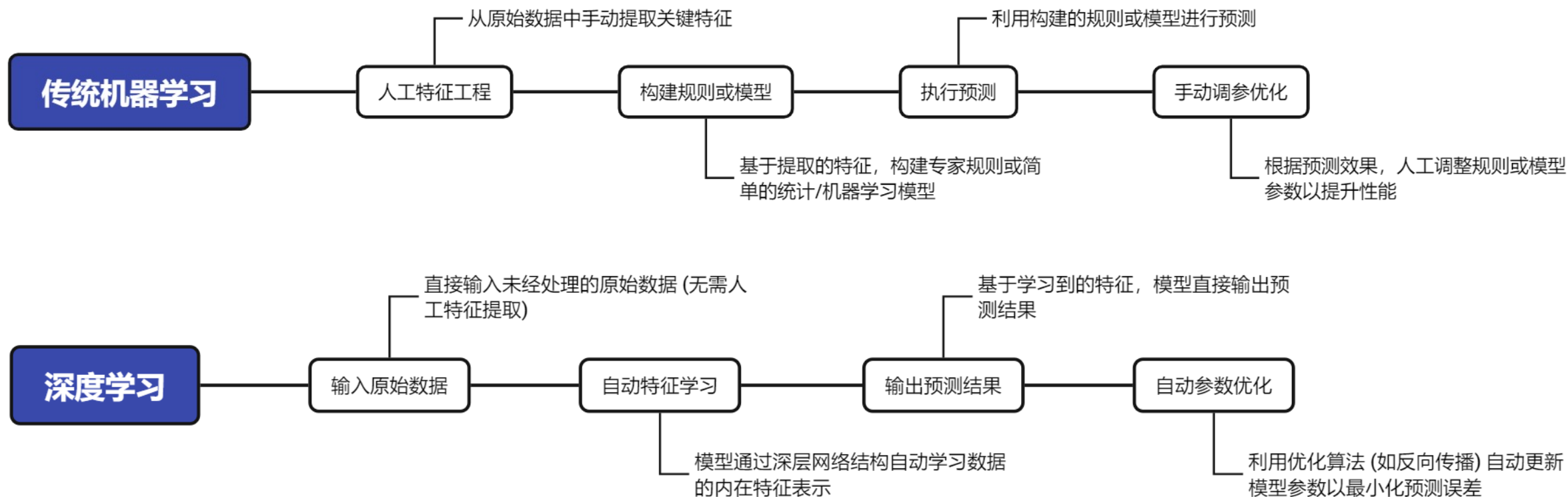


传统机器学习与 深度学习对比

传统机器学习与深度学习对比



训练流程对比



传统机器学习优点

- 原理透明性强：规则逻辑清晰，预测过程可解释
- 实现复杂度低：对计算资源要求低，部署成本小
- 适用性明确：在小规模、静态场景中高效稳定

传统机器学习缺点

- 强领域知识依赖：需专家设计特征与规则，泛化能力弱
- 动态适应性差：难以处理环境突变或非线性关系
- 精度瓶颈显著：对复杂模式识别能力有限

深度学习优点

- 端到端特征学习：自动挖掘数据内在模式，减少人工干预
- 高维建模能力强：可处理图像、时序等复杂数据结构
- 预测精度与泛化性优：在动态场景中保持高准确性

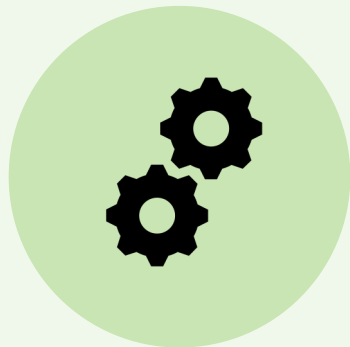
深度学习缺点

- 黑箱效应突出：决策逻辑不透明，可解释性差
- 训练成本高昂：需大量算力与存储资源，耗时长
- 数据依赖性高：模型效果受限于标注数据质量与规模

传统机器学习与深度学习对比



华南农业大学
South China Agricultural University



传统机器学习

规则驱动
静态场景
小规模数据
解释性强



深度学习

数据驱动
动态环境
大规模数据
预测性能优秀

课间视频



华南农业大学
South China Agricultural University



[新闻直播间]第十七届国际交通技术与设备展览会 人工智能等新技术在交通领域不断推广应用



特征工程：人教机器理解数据

本质：人工设计数据转换规则，特征设计决定模型上限
经典应用：

- 图像领域：边缘检测器 | 纹理描述符 | 颜色直方图
- 文本领域：关键词提取 | 词频统计 | TF-IDF

优势：

- 过程透明可解释
- 模型轻量易部署

局限：

- 依赖专家经验
- 难以捕捉非线性特征
- 跨场景迁移成本高



表征学习：机器自主理解数据

本质：端到端多层次特征抽象，数据质量决定模型潜力
实现路径：

- 浅层网络：提取基础特征（边缘/颜色）
- 中层网络：组合局部模式（纹理/结构）
- 深层网络：构建语义概念（对象/意图）

优势：

- 自动挖掘复杂模式
- 支持跨模态统一处理
- 强大迁移能力（预训练+微调）

应用突破：

- 图像识别
- 语音处理
- 自然语言理解



●———— 常见深度学习模型介绍 ————●

常见深度学习模型介绍



华南农业大学
South China Agricultural University

卷积神经网络（CNN）

模拟人类视觉皮层分级处理机制

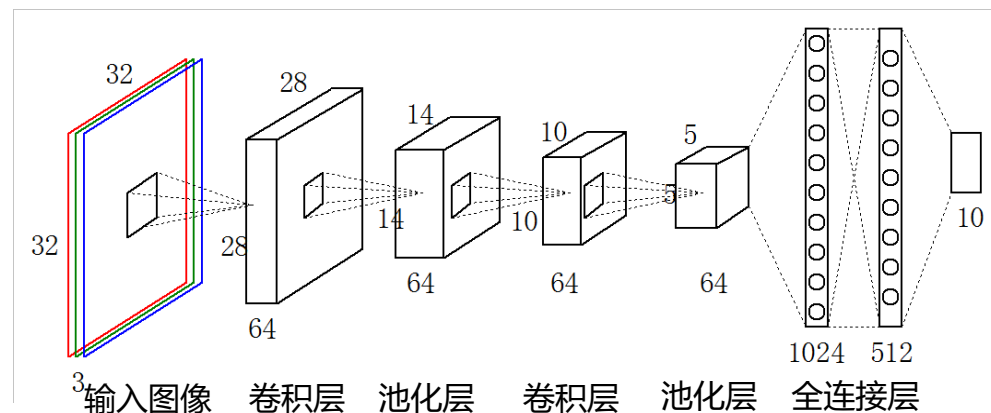
工作方式：逐层提取图像特征，最终映射为预测结果。

核心组件：

- 卷积层：局部特征提取（滤波器滑动计算）
- 池化层：特征降维与平移不变性
- 全连接层：全局信息整合与分类

相比传统方法，CNN 的优势在于不依赖人工设计特征，能自动从原始图像中学习有效表达。

CNN 广泛应用于图像识别、人脸检测和医学影像分析等领域。



LeNet-5模型结构

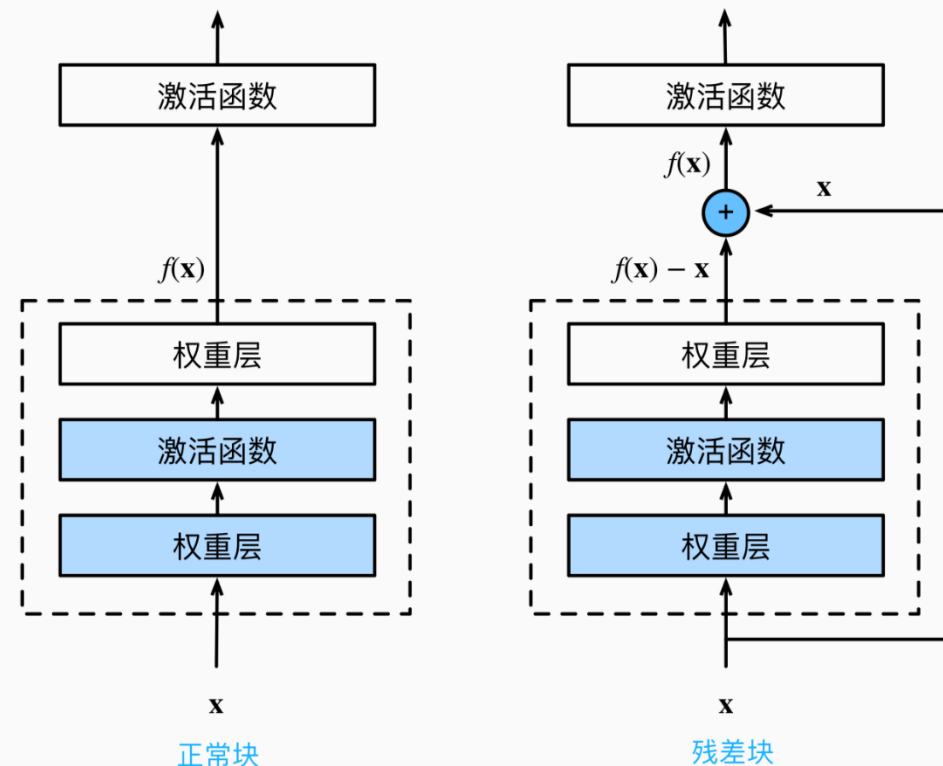
常见深度学习模型介绍



残差网络 (ResNet)

ResNet 是 CNN 的重要改进，由微软亚洲研究院何恺明等研究人员在 2015 年提出。

核心思想是引入残差连接 (Residual Connection)，通过增加恒等映射通路，实现信息在网络中的跳跃式传播。



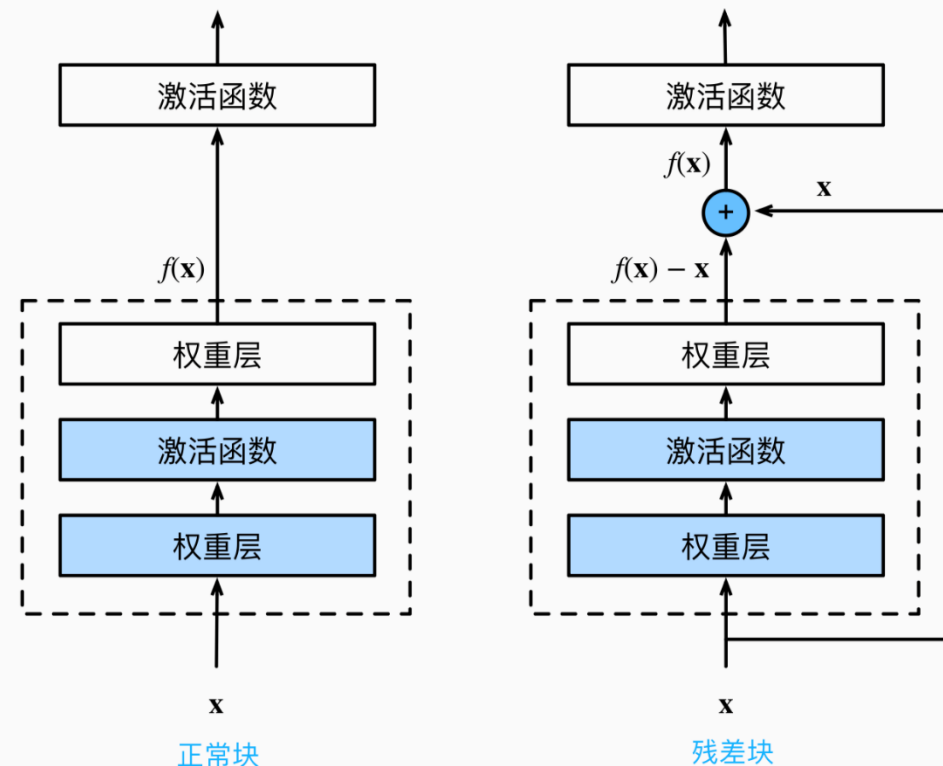
常见深度学习模型介绍



残差网络 (ResNet)

ResNet 解决了深层网络的梯度消失和性能退化问题，使训练上百层甚至上千层的深度神经网络成为可能。

ResNet 不仅提升了图像识别的准确率，还奠定了后续更复杂神经网络（如 ResNeXt 和 DenseNet）的基础。



常见深度学习模型介绍

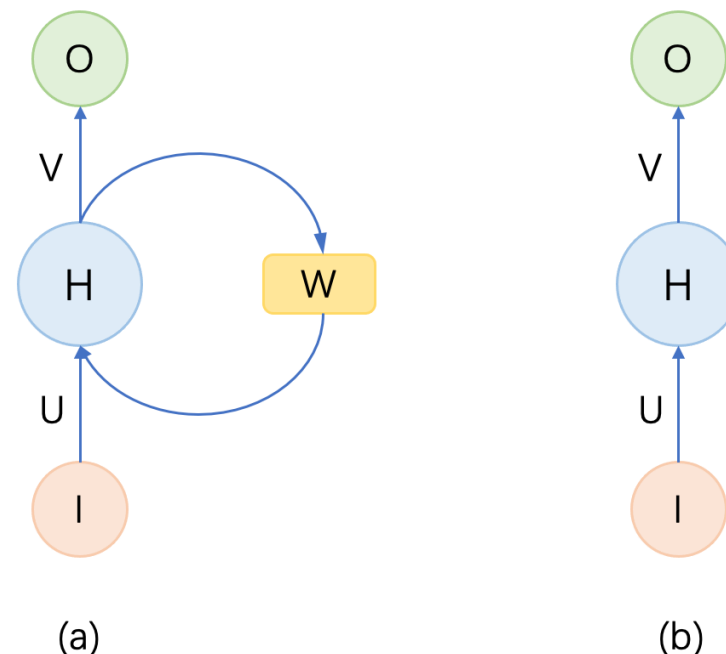


循环神经网络（RNN）

RNN 是专为处理序列数据（如时间序列、语言）设计的模型。

RNN具有“记忆”机制：每一步的输出不仅取决于当前输入，还受上一步状态的影响。

但RNN易受梯度消失问题影响，导致难以学习长期依赖关系



图(a)是一个典型的RNN结构图。

图(b)是一个普通的前馈神经网络。

RNN相比于一般前馈网络，增加了一个保存上下文信息的权重矩阵 W ，也即每次计算输出不仅要考虑当前输入数据，还要考虑序列数据的上下文信息。

常见深度学习模型介绍

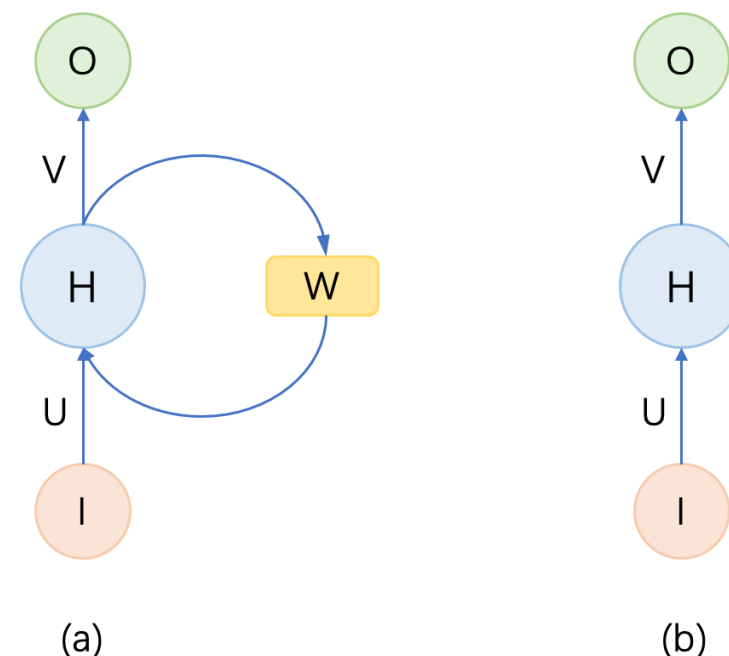


华南农业大学
South China Agricultural University

循环神经网络（RNN）

改进方案：实际应用中常通过长短期记忆网络（LSTM）或 Transformer 等模型解决上述问题。

应用现状：Transformer 等模型已取代 RNN 在多数任务中的主导地位。RNN 在小规模、低资源的时间序列场景中仍有独特价值（如简单时序预测、轻量级应用）。



图(a)是一个典型的RNN结构图。
图(b)是一个普通的前馈神经网络。
RNN相比于一般前馈网络，增加了一个保存上下文信息的权重矩阵W，也即每次计算输出不仅要考虑当前输入数据，还要考虑序列数据的上下文信息。

长短期记忆网络（LSTM）

LSTM 是专门为解决 RNN 长序列建模中的梯度消失问题而设计的改进型循环神经网络。

LSTM创新的引入三大门控结构：

- 输入门：控制新信息的写入
- 遗忘门：决定旧信息的保留程度
- 输出门：调节信息的输出

长短期记忆网络（LSTM）

核心优势：有效解决长期依赖问题，实现对长时序信息的记忆能力。

在语音识别、机器翻译、情感分析等任务中表现突出，尤其擅长捕捉序列中的时序特征。

扩展能力：

- 训练增强技术：结合双向结构（BiLSTM）集成注意力机制
- 模型融合方向：与 CNN、Transformer 等模型融合以提升综合表达能力

Transformer

Transformer是一种基于自注意力（Self-attention）机制的深度神经网络结构，最初用于自然语言处理。Transformer彻底摒弃 RNN 的序列计算方式，采用多头自注意力（Multi-head Self-attention）机制，并行处理整个输入序列，实现高效的信息交互。

核心架构：

每层包含四大组件：多头自注意力、前馈神经网络、层归一化、残差连接

具备良好的扩展性与训练稳定性

Transformer已成为大语言模型的核心架构（如 BERT、GPT 系列、DeepSeek 系列）。在文本生成、问答系统、代码生成等任务中表现卓越。并且已经扩展应用至图像、语音等跨模态场景。



思考与讨论

哪些模型可以应用到“交通拥堵预测”案例中，说明各模型适用场景。



●———— 常见深度学习框架介绍 ————●

常见深度学习框架介绍



华南农业大学
South China Agricultural University

深度学习框架

深度学习框架作为深度学习模型的“工具箱”，为机器学习的进行提供从设计到部署的全流程支持，解决了“仅有模型结构无法落地”的痛点，是连接理论与实际应用的桥梁。

深度学习框架的核心功能：

- 自动微分：自动计算梯度（如反向传播）
- 优化器集成：SGD、Adam 等训练算法开箱即用
- 数据管道：高效数据加载与预处理（如批处理、增强）
- 硬件加速：GPU/TPU 并行训练支持
- 可视化监控：训练过程实时跟踪（如损失曲线）
- 部署工具链：模型压缩、跨平台导出（移动端/边缘设备）

模型是建筑设计图，
框架是施工工具与平台



深度学习模型

从设计到实现



深度学习框架



TensorFlow 是由 Google Brain 团队开发的开源深度学习框架，支持从桌面训练到云端部署的完整流程，并通过计算图（Computation Graph）机制优化计算性能。专注生产环境的高性能部署，为工业级 AI 应用提供全流程支持，是工业级深度学习框架的标杆。

常见深度学习框架介绍



华南农业大学
South China Agricultural University



核心特性:

- 计算图 (Computation Graph) 机制: 通过图优化显著提升计算性能
- 全流程支持: 覆盖从桌面训练到云端部署的完整 workflow

技术优势:

- 灵活的工具库: 支持构建/训练各类深度学习模型
- 高性能扩展能力: 分布式训练、TPU 硬件加速
- 跨平台部署: 移动端 - TensorFlow Lite、Web端 - TensorFlow.js



PyTorch 是由 Meta 人工智能研究院（原名Facebook AI Research）开发的开源框架，提供灵活的调试与操作方式，适合开发复杂的网络结构和新颖算法。其社区活跃，文档清晰，支持如 Hugging Face 的 Transformers 库等热门工具库。近年来，PyTorch 也不断向工业界的生产部署靠拢，已成为学术与工业兼顾的主力框架。

常见深度学习框架简介



华南农业大学
South China Agricultural University



核心特性:

- 动态计算图: 提供灵活的调试与操作方式
- 开发友好性: 特别适合构建复杂网络结构和实现新颖算法

生态优势:

- 活跃开发者社区与清晰文档体系
- 深度集成热门工具库 (如 Hugging Face Transformers)



MXNet 是由亚马逊公司（Amazon）支持的开源框架，以其高效的计算性能和可扩展性闻名。MXNet 支持多种编程语言，适合大规模分布式训练。MXNet 是 AWS（Amazon Web Services）上原生支持的深度学习框架之一，适用于大规模并行训练场景，尤其适合构建需要云端计算资源的大型模型系统。但近年来其社区活跃度有所下降，已逐渐被TensorFlow 和 PyTorch取代。

常见深度学习框架介绍



华南农业大学
South China Agricultural University



核心优势:

- 高效计算性能与强可扩展性
- 多语言支持 (Python, C++, R, Scala 等)
- 专为大规模分布式训练设计

云集成能力:

- AWS (Amazon Web Services) 原生支持框架
- 特别适合云端大规模并行训练场景
- 优化构建依赖云端资源的大型模型系统

常见深度学习框架介绍



华南农业大学
South China Agricultural University



PaddlePaddle（飞桨）是由百度开发的开源深度学习平台，支持从研究到工业应用的全流程。在自然语言处理和计算机视觉等领域有丰富的工具和模型库，具备良好的工程化能力和产业适配性，尤其在中文自然语言处理、语音识别、图像识别等任务中有优异表现。

领域优势：

- 自然语言处理（NLP） 和 计算机视觉（CV） 领域领先
- 提供丰富的预训练模型库和专用工具
- 在中文NLP、语音识别、图像识别任务中表现优异

常见深度学习框架介绍



华南农业大学
South China Agricultural University



工程化能力：

- 强大的产业适配性和工程部署能力
- 针对国产硬件/操作系统的深度优化

本土化优势（vs TensorFlow/PyTorch）：

- 中文支持：原生中文文档和社区
- 文档本地化：完整的中文技术资料
- 国产化适配：自主可控的国产AI生态

[M]^s MindSpore

MindSpore 是华为推出的开源深度学习框架，支持端、边、云协同训练与部署，具有高效计算和隐私保护等特点。

核心特性：

- 全场景支持：端（设备）、边（边缘计算）、云协同训练与部署
- 高效计算：通过自动并行化技术优化资源利用率
- 隐私保护：内置安全机制保障数据安全

核心技术：

- 自动并行化：智能切分模型实现分布式训练加速
- 图算融合：融合计算图优化，显著提升训练/推理效率

[M]^s MindSpore

应用场景：

- 高性能计算需求：大规模模型训练
- 隐私敏感场景：医疗、金融等数据安全领域
- 国产化替代：自主AI技术生态建设

硬件适配：

- 深度优化适配国产昇腾（Ascend）芯片
- 支持端边云全栈国产硬件生态

Jittor 计图

Jittor 是由清华大学胡事民院士团队开发的开源深度学习框架，以其高性能和易用性著称。Jittor 通过即时编译（Just-In-Time Compilation）技术，显著提升了模型训练和推理的速度，可支持动态计算图和自动微分，适合研究和工业应用。

核心技术：

- 即时编译（JIT）技术：动态优化计算图，显著提升运行效率
- 动态计算图：支持灵活构建和修改模型结构
- 自动微分：简化梯度计算流程

常见深度学习框架介绍



华南农业大学
South China Agricultural University



领域应用：

- 在计算机视觉（CV） 和自然语言处理（NLP） 领域广泛应用
- 特别适合计算密集型任务和快速实验验证场景

应用特性：

- 易用性强： 简洁API设计降低学习成本
- 灵活开发： 支持研究原型快速迭代
- 工业级适用： 满足生产环境部署需求

性能优势：

- 训练/推理速度显著提升（相比传统框架）
- 高效计算： 优化底层算子实现

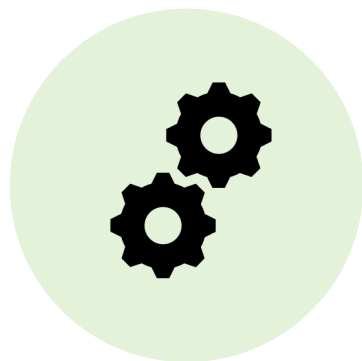
常见深度学习框架介绍



框架	开发公司	核心优势	适用场景	技术亮点
TensorFlow	Google	工业部署能力	生产环境/跨平台部署	计算图优化/TPU加速
PyTorch	Meta	灵活开发/强大生态	研究+工业落地	动态图优先/HuggingFace集成
MXNet	Amazon	分布式训练性能	AWS云平台大规模训练	多语言支持
PaddlePaddle	百度	中文任务优化/国产适配	中文NLP/CV/语音	本地化工具链
MindSpore	华为	端边云协同/国产芯片优化	自主生态/隐私计算	图算融合/昇腾芯片适配
Jittor	清华大学	JIT编译加速/动态图	高效计算研究	即时编译技术



课堂总结



传统机器学习

规则驱动

人工特征工程（耗时）
静态场景适用（小数据）
解释性强但性能有限



深度学习

数据驱动

自动表征学习（高效）
动态预测优化（大数据）
高精度黑盒模型

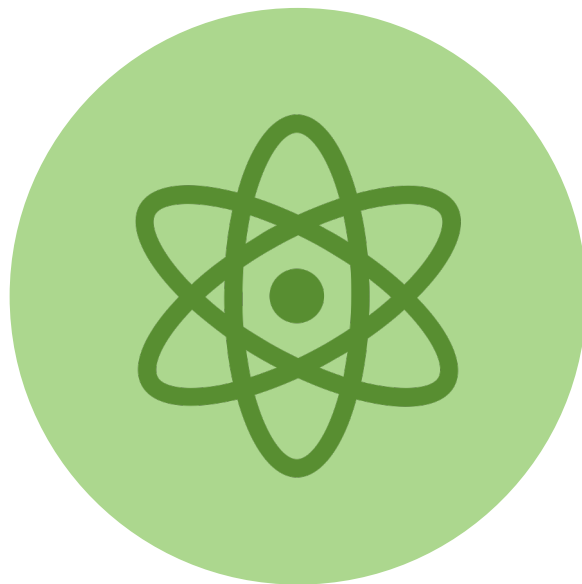
“当数据复杂、规则难寻，深度学习是必然选择”



回顾交通拥堵预测的案例。传统机器学习方法在该案例中遇到了哪些具体瓶颈？

深度学习的哪些核心特性（如自动特征学习、端到端建模、高维建模能力）能够有效解决这些瓶颈？

除了交通预测，你能想到生活中或你所知领域中的哪些问题，同样面临传统机器学习难以解决的复杂性和动态性挑战？



下载实验文件，根据实验引导体
验使用PyTorch搭建简易CNN模型完
成手写数字识别任务