

— AI深度学习 —

深度学习介绍

Introduction of Deep Learning



深度学习概述



认识数据集



神经网络训练工具

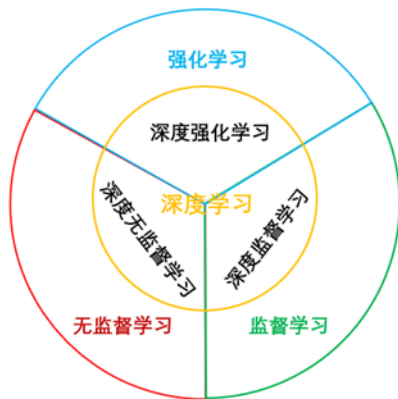
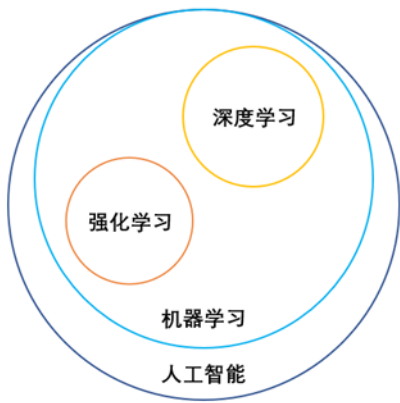
1.深度学习概述

深度学习简介_概念关系

人工智能：研究使计算机来**模拟人的某些思维过程**和**智能行为**（如学习、推理、思考、规划等）的学科

机器学习：从**数据**中进行**学习**，提高计算机程序效果的学科

深度学习：机器学习领域的一个分支，以**人工神经网络**为架构，对**数据**进行表征学习的方法



易错概念

机器人 ≠ 人工智能

大数据 ≠ 人工智能

人工智能 ≠ 深度学习

机器学习 ≠ 深度学习

图像识别 ≠ 深度学习

机器翻译 ≠ 深度学习

人工智能的应用场景：计算机视觉（看），自然语言处理（读写），大数据（数据处理），机器人（行动）

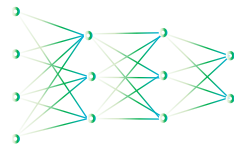
深度学习简介_与传统机器学习的区别

面试问题1：深度学习相比传统机器学习的主要不同在是什么？



先进行特征工程：

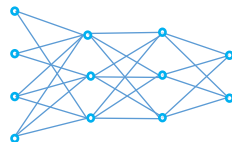
根据人类经验设计要提取的特征，由人类自己标注这些特征。（如：萼片长度，萼片宽度。）



传统机器学习：无法自己提取特征

鸢尾花分类模型A

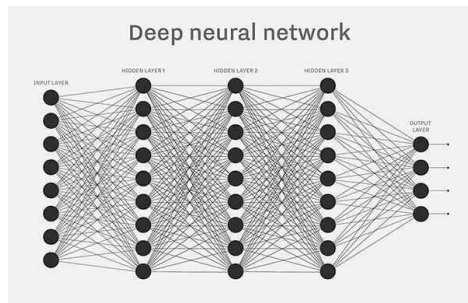
面试问题2：深度学习相比传统机器学习的优越性体现在什么地方？



深度学习：网络自行提取出不同层次的特征

鸢尾花分类模型B

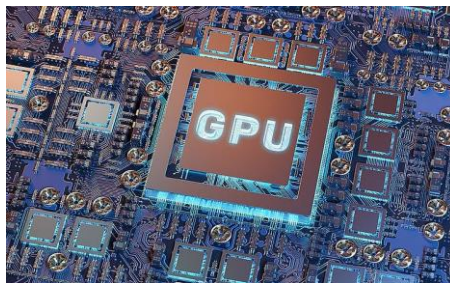
深度学习的三大关键要素



深度神经网络模型



大规模数据集

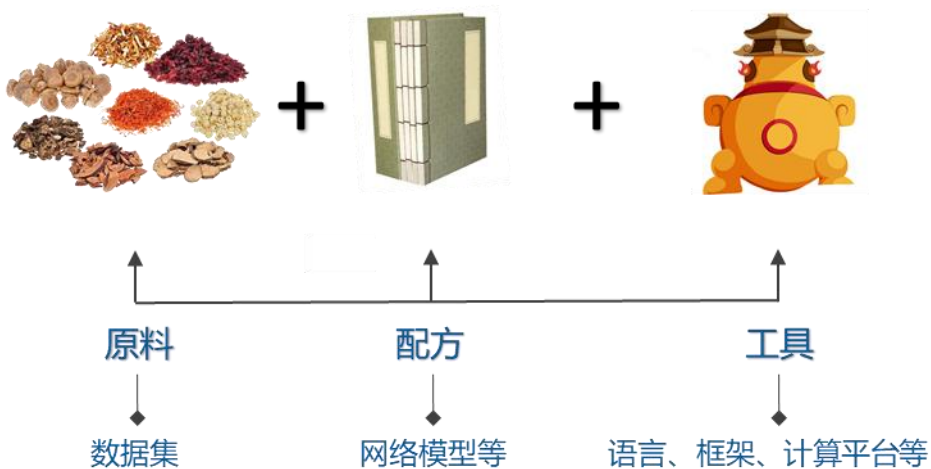


大规模运算资源 (如GPU)

面试问题1：深度学习的三大关键要素是什么？

面试问题2：神经网络在几十年前就出现了，为什么深度学习最近几年才火起来？

深度学习的三大关键要素

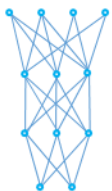


深度学习简介

大规模数据集

运算平台

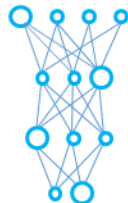
未训练的
神经网络模型



训练

从**现有数据**中学习经验

训练已有模型
获得新的经验



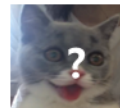
猫

狗



推理

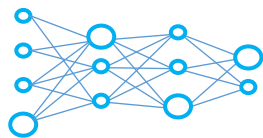
将此经验应用于**新数据**



深度学习的应用

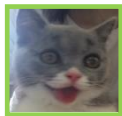


输入

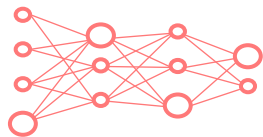


输出

5



输入

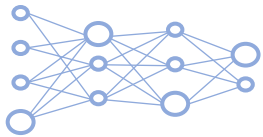


输出

cat



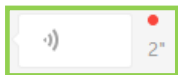
输入



输出

love

图像识别



输入



输出

Hello world

机器翻译

语音识别

深度学习的应用

自然语言处理

- 机器翻译
- 文本分类
- 文本生成
- 问答系统

机器人学

- 决策控制
- 路径规划/导航
- 多智能体/博弈

让机器像人
一样读写

让机器像人
一样看懂

深度学习

让机器像人
一样运动

让机器像人
一样听懂

计算机视觉

- 图像识别
- 目标检测
- 图像分割
- 人脸检测
- 人体姿态识别
- 图像合成

智能语音

- 语音识别
- 语音处理
- 语音合成

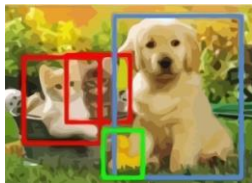
深度学习的应用

计算机视觉任务举例



→ 猫

图像识别



→ 猫
狗
鸭子

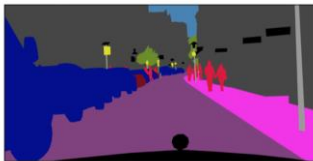
目标检测



人脸关键点检测



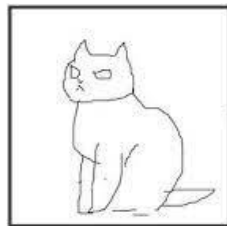
(a) Scene image



(b) Groundtruth

■ road ■ sidewalk ■ building ■ wall ■ fence ■ pole ■ traffic light ■ traffic sign ■ vegetation ■ terrain
■ sky ■ person ■ rider ■ car ■ truck ■ bus ■ train ■ motorcycle ■ bicycle

图像分割



pix2pix
process



图像生成

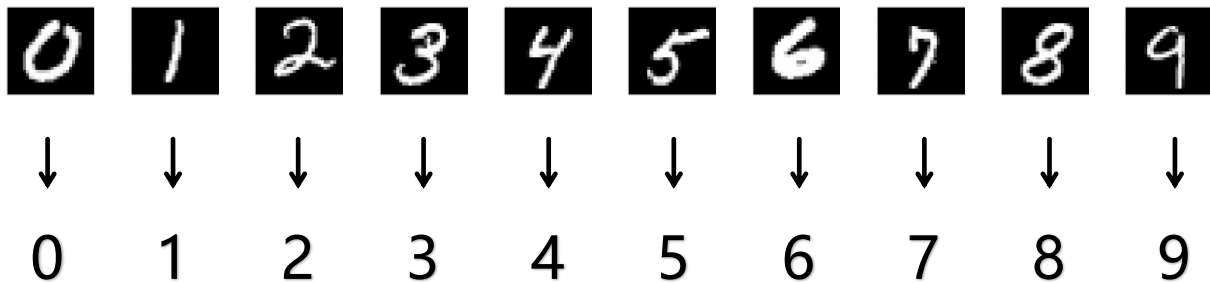
2.认识数据集

数据集定义

数据集 (dataset) 是指很多样本组成的集合

图像识别：给一张图片，判断图片的类别

图像识别数据集的组成：图片+标签



经典数据集-MNIST

MNIST

[Yann LeCun](#), Courant Institute, NYU
[Corinna Cortes](#), Google Labs, New York
[Christopher J.C. Burges](#), Microsoft Research, Redmond



“NIST” 代表国家标准和技术研究所 (National Institute of Standards and Technology)，是初收集这些数据的机构。M代表 “修改的 (Modified)”。

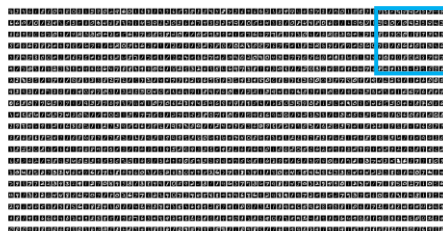
为了便于大家使用， MNIST中的数据全都进行了预处理：保证尺寸相同、图像中的字体居中等。

对于希望在真实数据上实践深度学习的人群， **MNIST** 是一个很好的数据库。

经典数据集-MNIST

该数据集是7万张黑白手写字
体图片的集合

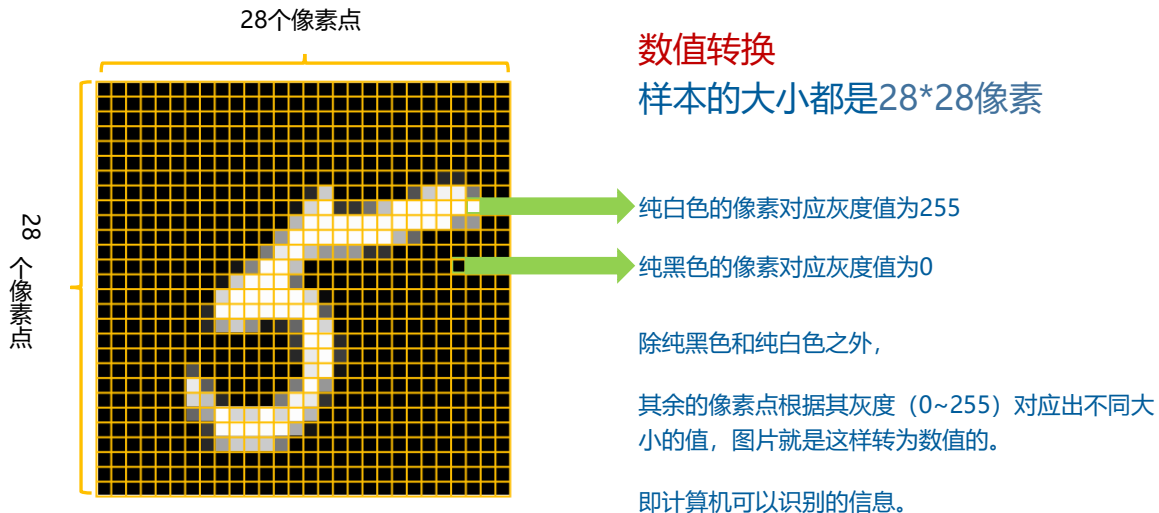
每张手写字体图片
都是相同尺寸



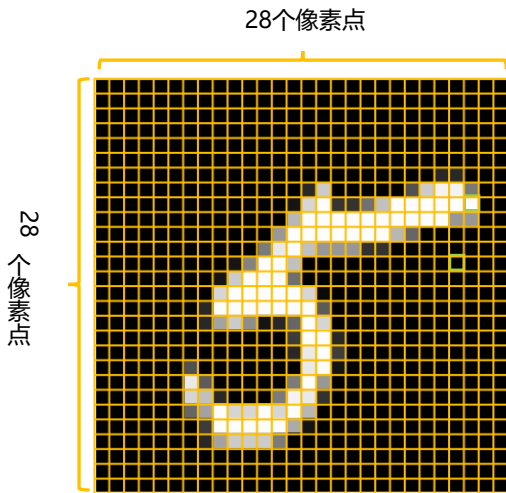
7	5	8	9	9	4	9	5	3	7	9
9	9	0	1	0	4	2	9	9	4	3
8	8	9	0	0	1	8	7	3	9	9
7	3	1	0	6	6	0	9	3	4	6
9	0	9	0	8	0	2	8	7	8	7
8	4	9	3	6	5	2	7	5	8	6



使用数据集-数值转换



使用数据集-Normalize



三个重要参数之 Normalize

面试问题：对输入数据进行归一化/标准化的目的和好处是什么？

1、Normalize用于做归一化

归一化前：784个像素将对应784个的灰度值，取值范围为0~255。

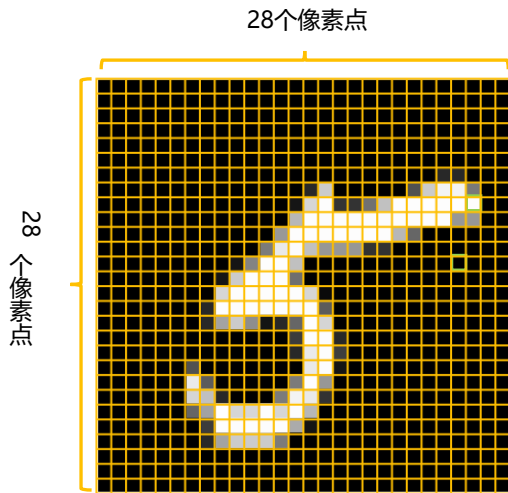
归一化后：784个像素将根据其灰度值大小转换为介于0~1之间的数值，一共有784个介于0~1之间的数值。

公式：原灰度值/255=新值，如纯白色灰度值为255，则新值为255/255=1。

拓展延伸：标准化，把数值范围变换到 [-1,1]

注意：归一化和标准化不会改变原始数据的数值排序。

使用数据集-flatten



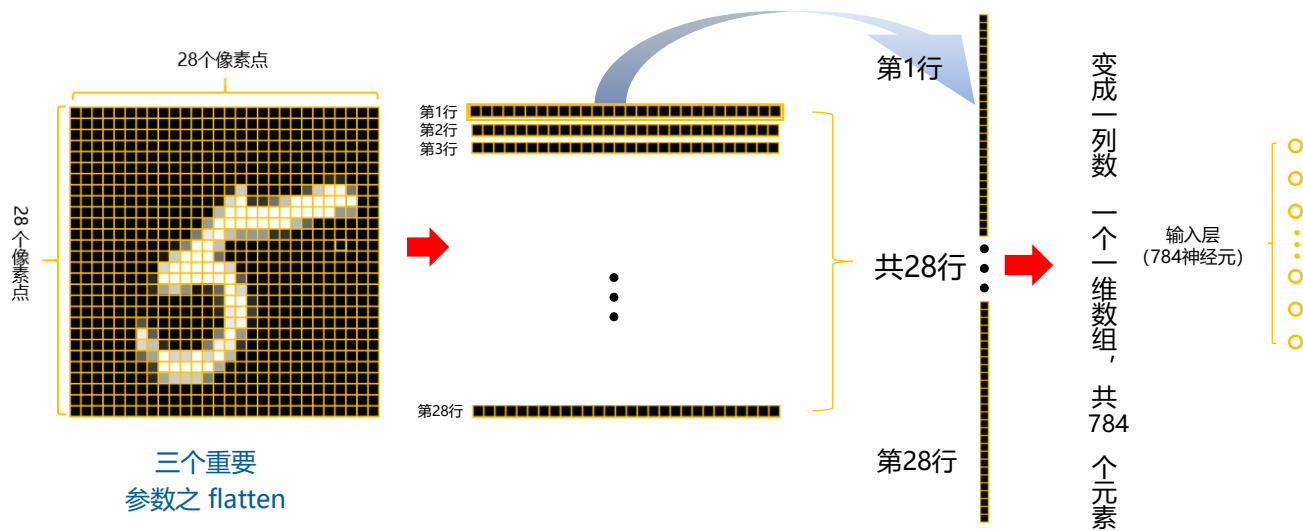
三个重要参数之 flatten

1、flatten用于设置是否展开输入图像数组

flatten=False: 计算机读取到的输入图像为一个
28*28的二维数组。

flatten=True: 计算机读取到的输入图像是由784
个元素构成的一维数组。（如下页所示）

使用数据集-flatten





使用数据集-one-hot


三个重要参数之 one_hot


数据集中的七万张手写字体样本，对应着七万个标签


One-Hot编码，又称为一位有效编码，主要是采用N位状态寄存器来对N个状态进行编码，每个状态都有他独立的寄存器位，并且在任意时候只有一位有效。


 → 0 → [1,0,0,0,0,0,0,0,0,0]


 → 1 → [0,1,0,0,0,0,0,0,0,0]


 → 2 → [0,0,1,0,0,0,0,0,0,0]


 → 3 → [0,0,0,1,0,0,0,0,0,0]


 → 4 → [0,0,0,0,1,0,0,0,0,0]

 → 5 → [0,0,0,0,0,1,0,0,0,0]

 → 6 → [0,0,0,0,0,0,1,0,0,0]

 → 7 → [0,0,0,0,0,0,0,1,0,0]

 → 8 → [0,0,0,0,0,0,0,0,1,0]

 → 9 → [0,0,0,0,0,0,0,0,0,1]

使用数据集-one-hot



标签【6】变为【0,0,0,0,0,0,1,0,0,0】

三个重要参数之 one_hot (独热编码)

one_hot 用于设置**标签**的表示

one_hot_label=False: 标签将**保存为【6】**

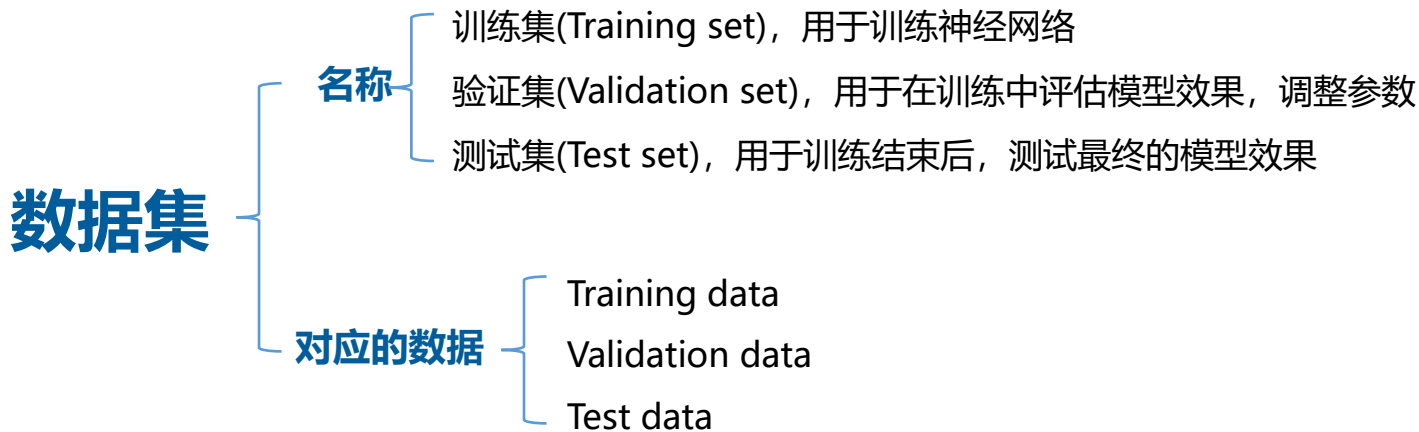
one_hot_label=true: 标签将**保存为one_hot表示**。

“one-hot表示” 即:

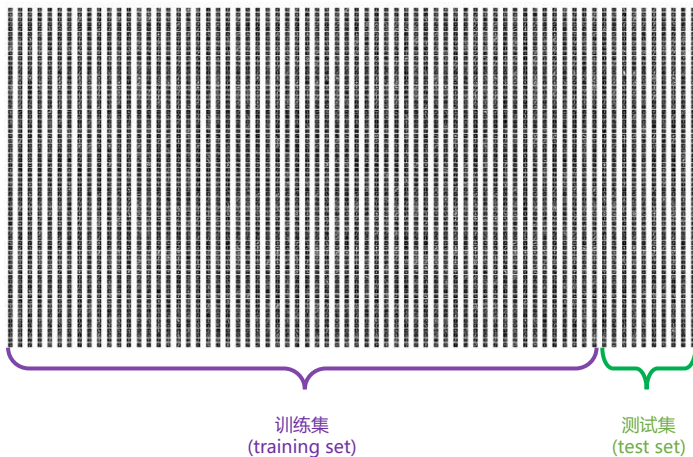
将标签由单一数值 **【6】**，表示为一维向量 **【0,0,0,0,0,1,0,0,0】**，此处标签的编码长度等于输出类别的数量：10。（如图左所示）

面试问题：为什么使用one-hot作为标签？优势是什么？

训练集、测试集与验证集



训练集、测试集与验证集

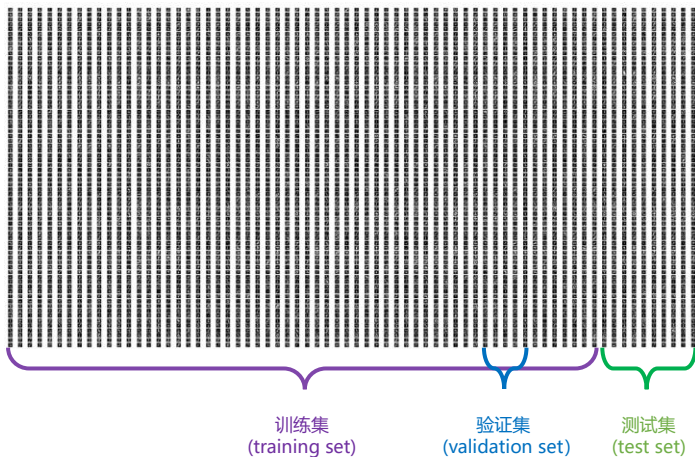


MNIST手写数字数据库:

训练集为60,000个样本

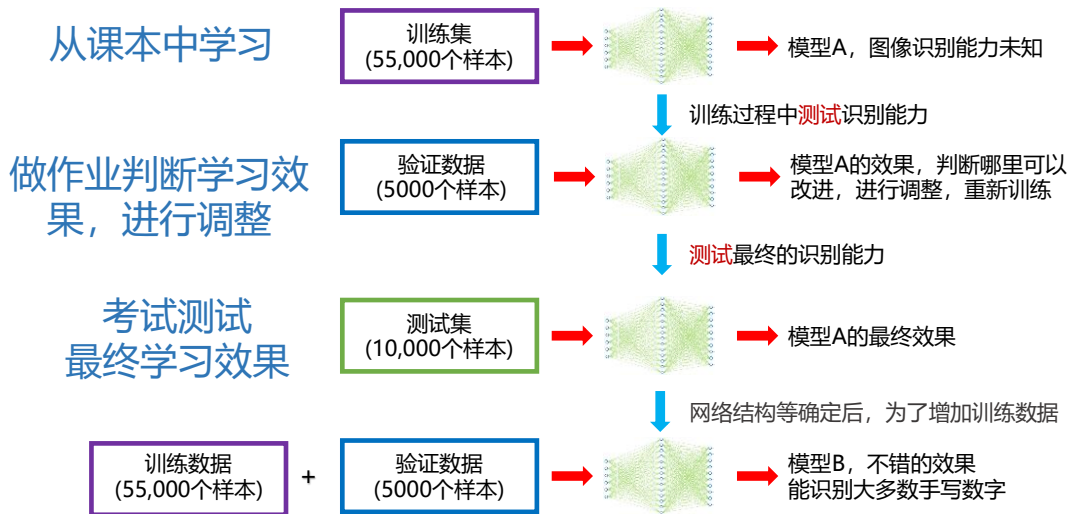
测试集为10,000个样本

训练集、测试集与验证集



训练集为60,000个样本，
测试集为10,000个样本，
也可在此基础上添加部分验证集
(比如5000个样本)，用来避免
过拟合。

训练集、测试集与验证集

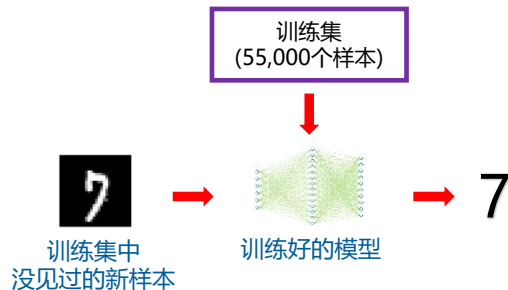


泛化与过拟合

泛化 generalization

机器学习的主要挑战：算法能够在**先前未观测到新输入上**表现良好，而不只是在训练集上表现良好。

在先前未观测到的输入上表现良好的能力被称为泛化能力（generalization）。



面试问题1：什么
叫模型的泛化能力？

面试问题2：数据
集中样本分布的基
本假设是什么？

面试问题3：为什
么数据集需要独立
同分布？

泛化与过拟合

泛化 generalization

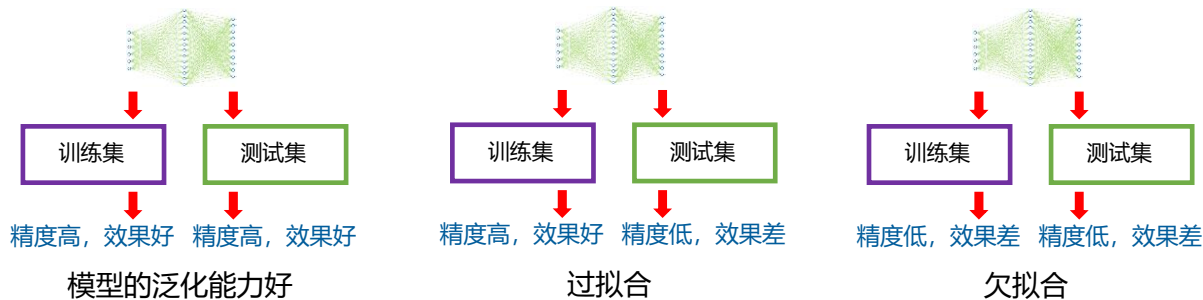
泛化误差 = 在训练集上的误差平均值 - 在测试集上的误差平均值

泛化误差可以反映模型的泛化能力，泛化误差越小，泛化能力越好

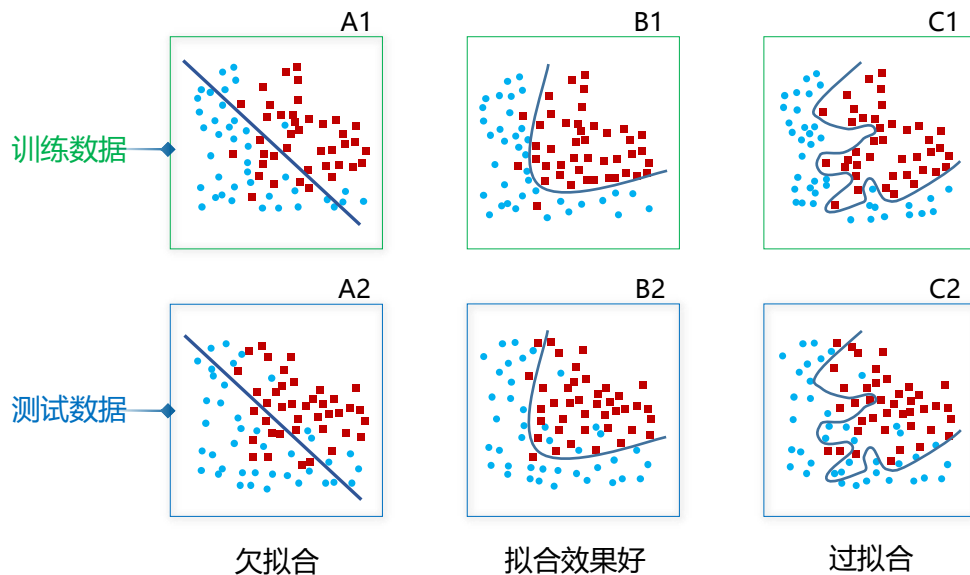
过拟合：泛化误差大，训练集上效果好，测试集上效果差（相当于对书本死记硬背，考试成绩差）

一般是数据集过小，模型过于复杂导致

欠拟合：训练集和测试集上效果都差，一般是模型复杂度不够导致



泛化与过拟合



面试题1：什么是过拟合？什么是欠拟合？

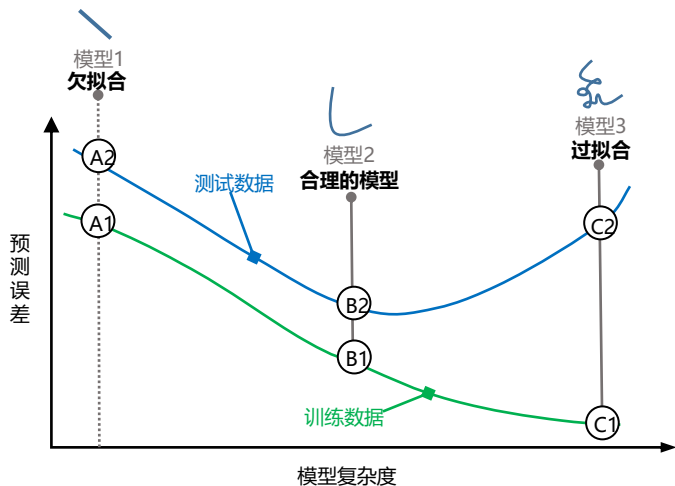
面试题2：在实验中，怎么判断是否发生了过拟合？

泛化与过拟合

拟合能力与模型复杂度的关系

过拟合：数据集过小，模型过于复杂导致（需要增大数据集，降低模型复杂度）

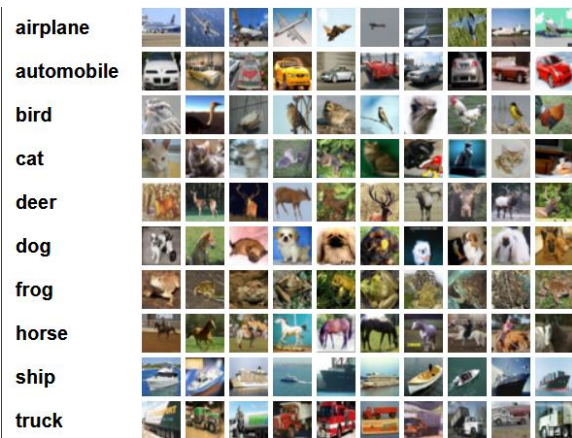
欠拟合：模型复杂度不够导致（需要增加模型复杂度）



面试问题1：过拟合和欠拟合产生的原因分别是什么？

面试问题2：如何解决过拟合或欠拟合的问题？

经典数据集-CIFAR-10



CIFAR：加拿大高等研究院（Canada Institute for Advanced Research）

CIFAR-10 和 CIFAR-100 由 Alex Krizhevsky, Vinod Nair 和 Geoffrey Hinton 收集。

CIFAR-10 数据集：

60000 个样本的集合（训练集 5 万、测试集 1 万）

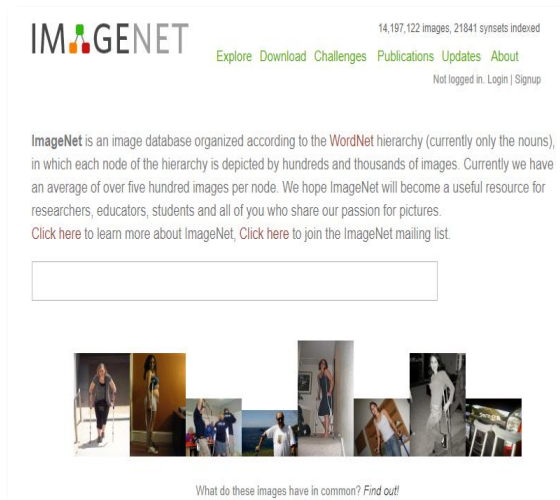
每个样本都是 32x32 像素的彩色图像

10 个类别：每个类别有 6000 个图像样本

左图：是 CIFAR-10 数据集集中的 10 个类，以及每个类中的 10 张随机图像样本。

CIFAR-100：该数据集类似于 CIFAR-10，不同之处在于它有 100 个类别。

经典数据集-ImageNet



ImageNet

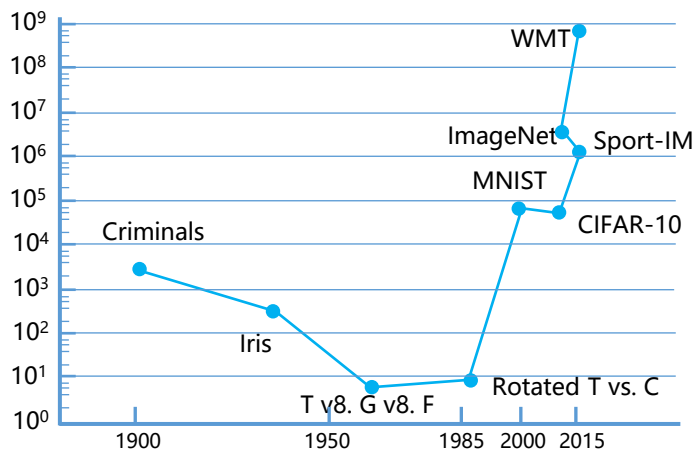


ImageNet是一个大型可视化数据库，是世界范围内图像识别领域最重要的数据库之一。

ImageNet包含了数千万个图像样本，并且这些样本被大家进行了手动标注。该数据集包含2万多个类别的图像数据。

2010年，ImageNet大规模视觉识别挑战赛（ImageNet Large Scale Visual Recognition Challenge, ILSVRC）启动，该大赛每年举办一次，不少优秀的神经网络模型，如AlexNet、GoogLeNet、VGGNet都是在该大赛中脱颖而出的。

数据集发展趋势



- 大数据时代下，数据量与日俱增
- 可用数据集的体量也越来越庞大
- 这些数据集作为能源驱动着深度学习的发展

3.神经网络训练工具

神经网络训练工具

深度学习对应的丹炉也就是训练模型需要的工具
如计算平台、深度学习框架、程序设计语言等

神经网络训练工具

处理器	CPU、GPU
工作站	英伟达DGX
云计算平台	亚马逊EC2、阿里云
软件库	TensorFlow、PyTorch、Caffe、MXNet
程序设计语言	C语言、Python

面试问题（最好写在简历上）：

你熟悉哪些编程语言？（如C, C++, Python）

你使用过哪些深度学习平台/软件库？（如果熟悉Pytorch、TF说明有实战经验）

你是否使用过GPU训练神经网络？（如果使用过GPU训练模型说明有实战经验）

你训练过哪些神经网络？（可能会根据训练过的网络提出相关的问题）