图像分类（计算机视觉的一个核心任务。）

1.输入：图片。输出：将图像分配到一组固定类别中的一个

2.面临问题：语义鸿沟（语义鸿沟指的是图像的底层视觉特征与高层次的语义理解之间的差距。计算机只能看到矩阵）

2.面临的挑战：

a. 视角变化(指的是同一物体或场景从不同角度拍摄所形成的图像之间的差异。)

b. 遮挡(指的是在图像或场景中，一个物体部分或全部被另一个物体遮挡。)

c. 类内变化(指的是同一类别内部个体之间的差异。)

d. 细粒度类别(指的是具有非常相似外观但属于不同子类别的物体。)

e. 背景杂乱(指的是图像背景中存在大量无关物体或元素，这些杂乱元素会干扰目标物体的检测和识别。)

f. 光照变化指的是同一场景或物体在光线条件改变时产生的视觉差异。

g. 变形指的是物体在三维空间中的形状发生改变，比如物体的扭曲、拉伸或压缩等。

3.作用：是其他任务的基石。如图像字幕生成，下围棋

机器学习：数据驱动方法

收集图像数据集和标签

使用机器学习训练分类器

在新图像上评估分类器

4.图像分类数据集：

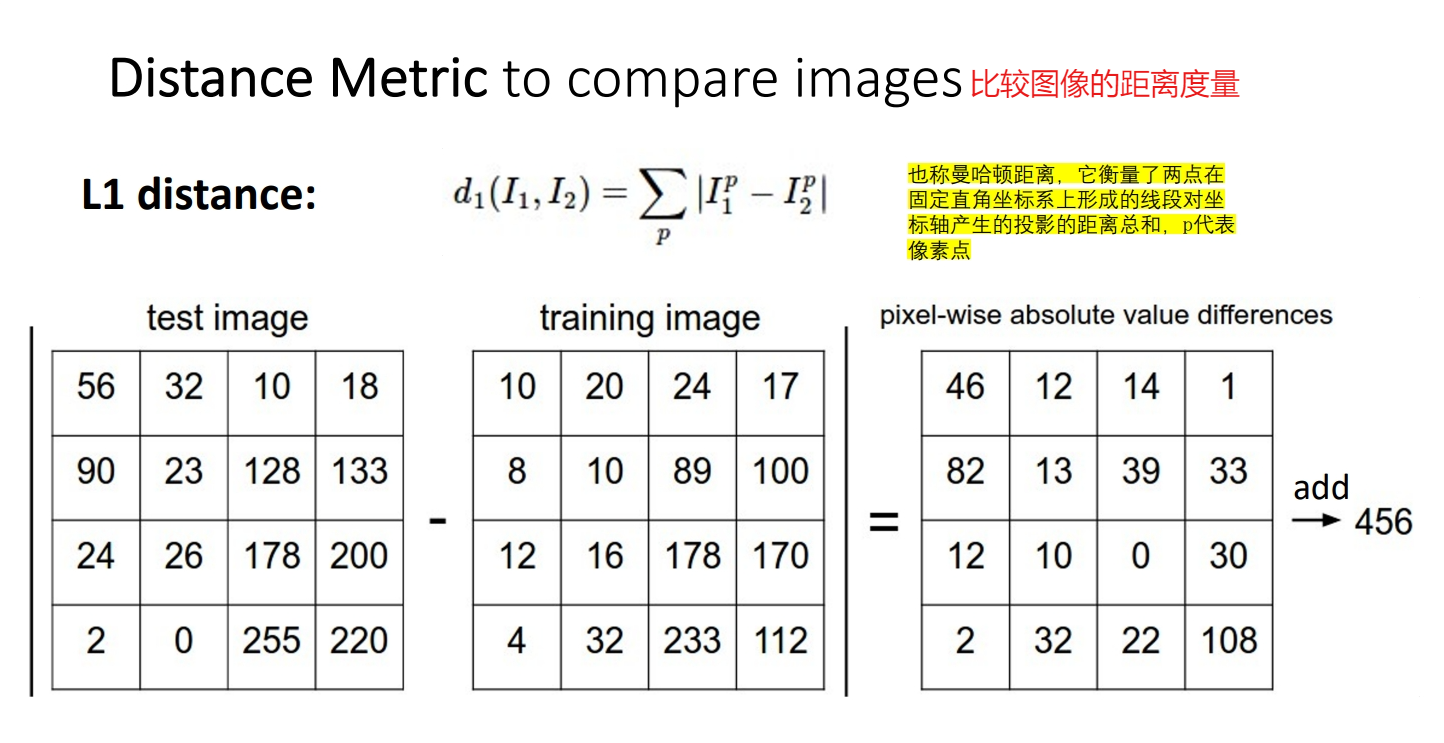
a.MNIST数据集的结果往往无法适用于更复杂的数据集！

b. CIFAR10

c. ImageNet

d. MIT Places

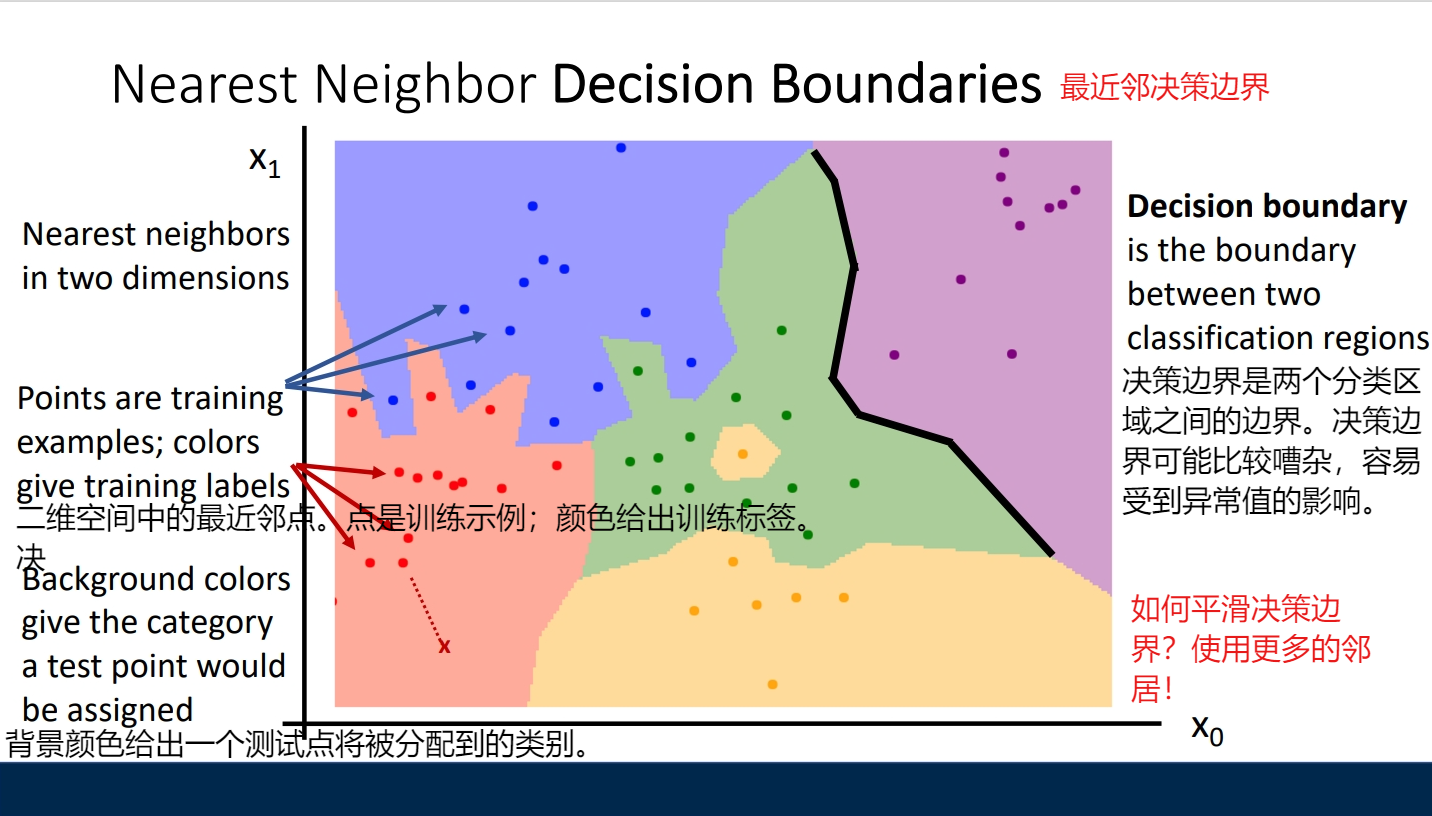
e. Omniglot



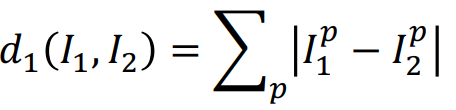
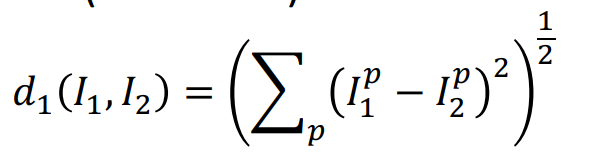


问题：有N个示例，训练速度如何？ 答案：O(1)

有N个示例，测试速度如何？ 答案：O(N)



K 最近邻：距离指标

L1（曼哈顿）距离L2（欧几里得）距离

如何设置超参数

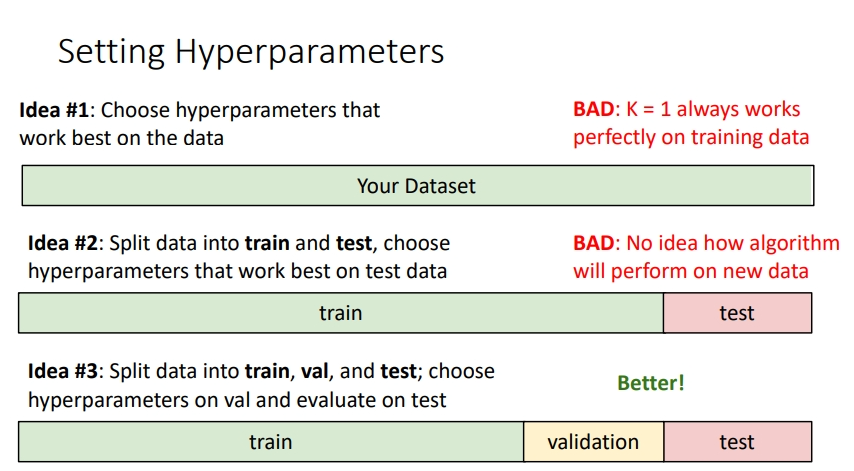
1：选择对数据最有效的超参数

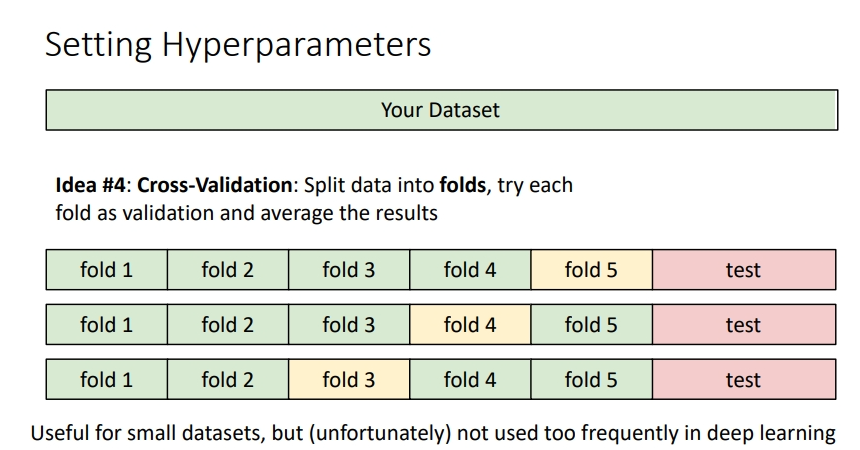
2：将数据拆分为训练和测试，选择最适合测试数据的超参数

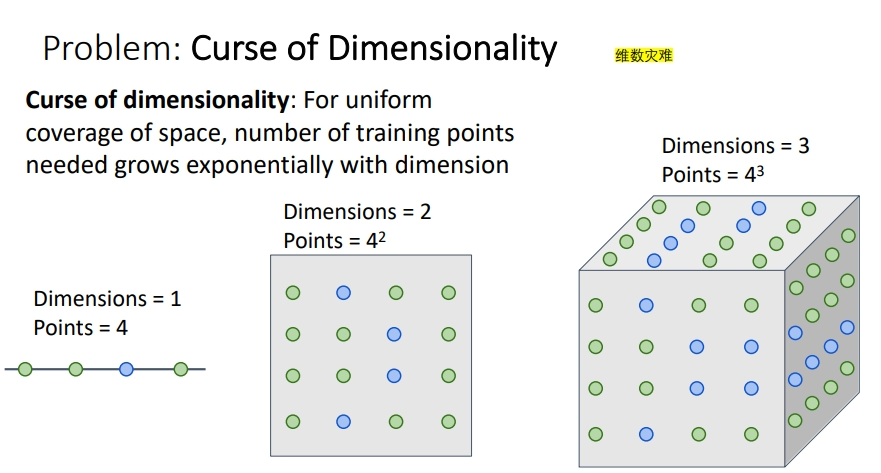
3：将数据拆分为 train、val 和 test;在 Val 上选择超参数并在测试时进行评估

4：交叉验证：将数据集分成几个折叠，每次使用一个折叠作为验证集，并平均每次的结果。

训练时面临的问题：维度灾难：为了在空间中均匀覆盖，所需的训练点数量随着维度的增加呈指数增长。







# 线性分类器

1.线性分类器为每个类别提供一个“模板”，对应于权重向量，用于将输入数据映射到该类别的得分。

2.分类器的任务是将图像中的每个像素与这些模板进行比较。如果某个像素与某个模板最相似，那么它就被分到相应的类别。

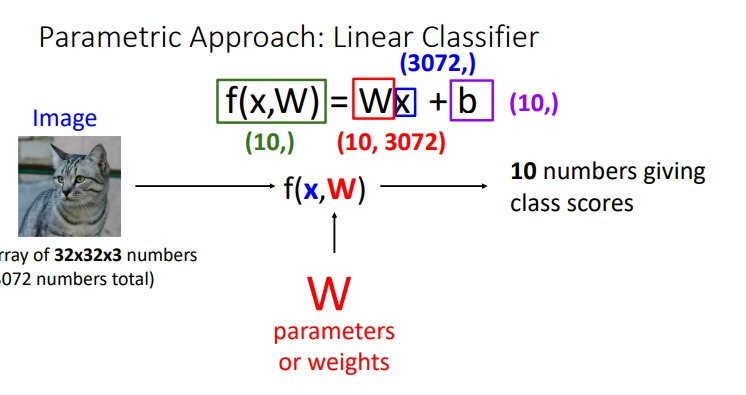
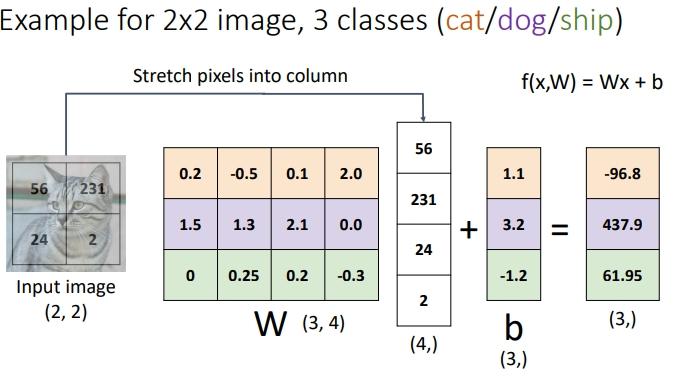
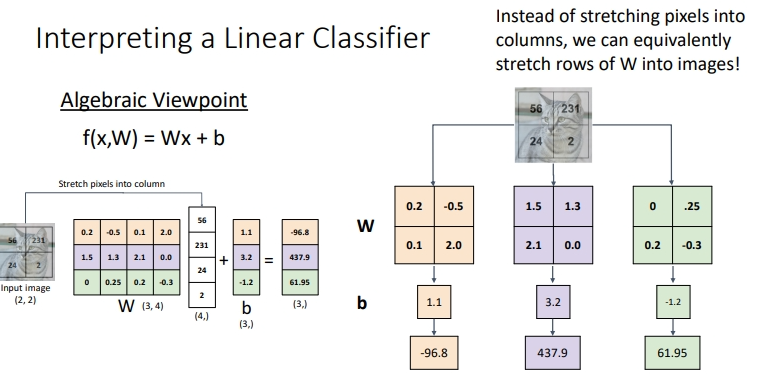
3.超平面分割了一个高维的空间

4.负损失函数有时被称为奖励函数，利润函数，效用函数，适应度函数

5.Kullback–Leibler（KL）散度，也称为相对熵或I-散度，是一种用于衡量 ，概率分布之间差异的统计距离

6.交叉熵 = 信息熵 + KL散度,由于信息熵一般为常数，因此说KL散,度等价于交叉熵。

计算预测值

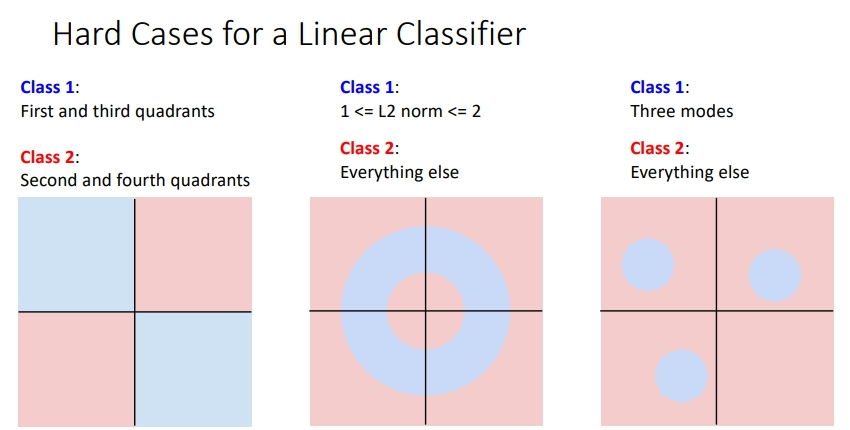
  

线性分类器的棘手点

对于类别1，数据分布在第一和第三象限，而类别2的数据分布在第二和第四象限。这种情况下，线性分类器很难找到一个好的分割超平面。

类别1的数据的L2范数在1到2之间，而类别2的数据则不受限制。这同样给线性分类器带来了困难。

类别1的数据有三个不同的分布模式，而类别2的数据则包含了除类别1外的所有其他情况。这种情况同样增加了线性分类器的分类难度。

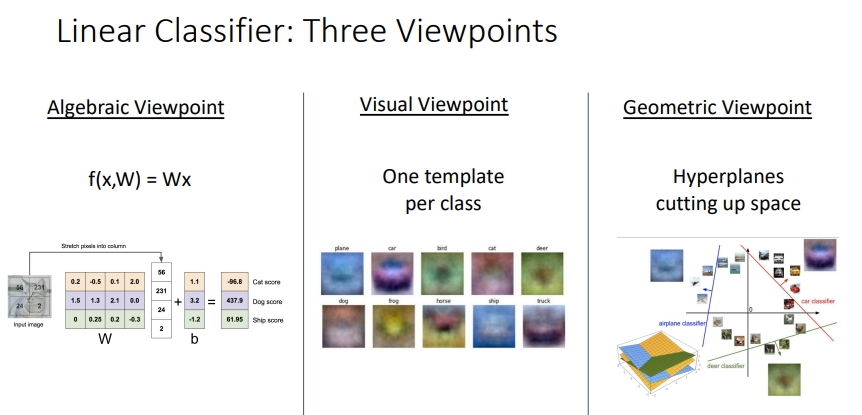


线性分类器的三个视角：

代数视角(Algebraic Viewpoint)：线性分类器的数学表达式为 f(x, W) = Wx，其中 W 是权重矩阵，x 是特征向量。这个视角着重从代数和数学表达式的角度来理解线性分类器。

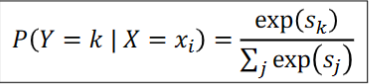
视觉视角(Visual Viewpoint)：线性分类器为每个类别提供一个“模板”或权重向量 W，用于将输入特征 x 映射到该类别的得分。这个视角着重从图像识别的角度来理解线性分类器。

几何视角(Geometric Viewpoint)：线性分类器通过将空间切割成不同的区域来划分不同类别。这些区域由超平面分隔，超平面的位置由权重 W 决定。这个视角着重从几何和空间划分的角度来理解线性分类器。



两种损失函数

1. 交叉熵损失函数

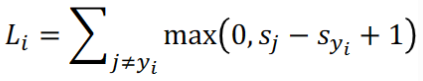


Li为单个数据样本的损失

Q: Li的最小/最大可能损失是多少？最小值为0，最大值为+∞。

Q：如果所有分数都是小随机值，损失是多少？A：-log(1/C)。

1. 多类SVM损失函数



. SVM 损失：在一个样本中，对于真实分类与其他每各个分类，如果真实分类所得的分数与其他各分类所得的分数差距大于或等于安全距离，则真实标签分类与该分类没有损失值；反之则需要计算真实分类与该分类的损失值； 真实分类与其他各分类的损失值的总和即为一个样本的损失值

①即真实标签分类所得分数大于等于该分类的分数+安全距离，S\_yi >=S\_j + △，那么损失值=0

②否则，损失值等于其他分类的分数 + 安全距离（阈值）- 真实标签分类所得的分数，即损失值=S\_j + △ - S\_yi

Q: 如果所有分数都是随机的，我们预计损失是多少？-log(1/C)。

Q:如果稍微改变最后一个数据点的分数，每个损失会发生什么？A: 交叉熵损失会改变；SVM损失保持不变。

Q: 如果我将正确类别的分数从10加倍到20，每个损失会发生什么？A: 交叉熵损失会减少，SVM损失仍为0。

