

小样本学习

1. 什么是相似度函数，相似度函数学习的基本思想是什么？

相似度函数（**Similarity Function**）是用于度量两个对象之间相似程度的函数。在机器学习和数据挖掘中，相似度函数通常用于比较两个样本、文档、图像或其他数据表示的对象之间的相似性。相似度函数的设计直接影响了模型的性能，因为它定义了模型对相似性的理解。在迁移学习中，相似度函数可以用于度量源领域和目标领域之间的相似性，或者在特定任务中度量不同任务之间的相似性。

基本思想是从大规模训练数据集中学习一个相似度函数（元学习），将相似度函数用于预测，通过比对查询样本与支持集每一个样本的相似度，找到得分最高的那个。

2. 请介绍相似度函数学习的流程，并解释孪生网络和三元组损失

(1) 在大规模数据集上训练一个孪生网络

(2) 给定一个 **k-way, n-shot** 的支持集

(3) 给定查询样本，预测其属于的分类

孪生网络是一种特殊的神经网络结构，有两个相同的子网络，它们共享相同的权重。这两个子网络分别处理输入样本，然后通过一个比较层（如欧氏距离或余弦相似度）将两个子网络的输出进行比较。孪生网络的训练目标是使得相似的样本在比较层中的距离较小，而不相似的样本的距离较大。对比损失是孪生网络中常用的损失函数。

三元组损失是一种用于训练相似度学习模型的损失函数，通常与三元组网络一起使用。在一个三元组中，包含一个锚点样本和两个样本，一个是正样本，另一个是负样本。三元组损失的目标是通过拉近锚点样本和正样本之间的距离，并推远锚点样本和负样本之间的距离，从而学到一个能够准确衡量相似性的模型。