

## 相似度函数篇

来自：AiGC面试宝典

宁静致远

2024年01月12日 06:35



扫码  
查看更

### • 相似度函数篇

- 一、除了cosin还有哪些算相似度的方法
- 二、了解对比学习嘛？
- 三、对比学习负样本是否重要？负样本构造成本过高应该怎么解决？

## 一、除了cosin还有哪些算相似度的方法

除了余弦相似度（cosine similarity）之外，常见的相似度计算方法还包括欧氏距离、曼哈顿距离、Jaccard相似度、皮尔逊相关系数等。

## 二、了解对比学习嘛？

对比学习是一种无监督学习方法，通过训练模型使得同一样本的表示更接近，不同样本的表示更远离，从而学习到更好的表示。对比学习通常使用对比损失函数，例如Siamese网络、Triplet网络等，用于学习数据之间的相似性和差异性。

## 三、对比学习负样本是否重要？负样本构造成本过高应该怎么解决？

对比学习中负样本的重要性取决于具体的任务和数据。负样本可以帮助模型学习到样本之间的区分度，从而提高模型的性能和泛化能力。然而，负样本的构造成本可能会较高，特别是在一些领域和任务中。

为了解决负样本构造成本过高的问题，可以考虑以下方法：

- 降低负样本的构造成本：通过设计更高效的负样本生成算法或采样策略，减少负样本的构造成本。例如，可以利用数据增强技术生成合成的负样本，或者使用近似采样方法选择与正样本相似但不相同的负样本。
- 确定关键负样本：根据具体任务的特点，可以重点关注一些关键的负样本，而不是对所有负样本进行详细的构造。这样可以降低构造成本，同时仍然能够有效训练模型。
- 迁移学习和预训练模型：利用预训练模型或迁移学习的方法，可以在其他领域或任务中利用已有的负样本构造成果，减少重复的负样本构造工作。