

NLP Trick 篇

来自：AiGC面试宝典

 宁静致远

2024年01月27日 19:14



扫码
查看更

- [NLP Trick 篇](#)
 - [一、怎么处理类别不平衡？](#)
 - [二、有了解其他模型去尝试解决长度限制的方案吗？](#)

一、怎么处理类别不平衡？

类别不平衡问题可以通过过采样、欠采样、生成新样本、集成学习等方法来解决。过采样方法包括随机过采样、SMOTE等；欠采样方法包括随机欠采样、Tomek Links等；生成新样本方法包括GAN、VAE等；集成学习方法包括Bagging、Boosting等。

二、有了解其他模型去尝试解决长度限制的方案吗？

Bert模型的长度限制问题主要是由于Transformer结构中的自注意力机制（self-attention mechanism）和位置嵌入（position embeddings）所导致的。这些机制使得Bert对于较长的序列处理非常耗时，并且占用大量的内存，从而限制了Bert在处理长序列任务上的性能。

为了解决这个问题，一些研究人员提出了一些改进型的模型，包括：

- Longformer：Longformer是一个基于Transformer结构的模型，它使用了一种新的自注意力机制，称为"Sliding Window Attention"，该机制可以在处理长序列时缓解Bert模型的计算和存储成本。
- Reformer：Reformer是一个基于哈希注意力（Hashing Attention）的Transformer模型，该模型可以有效地处理长序列，并且在一些NLP任务上表现良好。
- Performer：Performer是一种基于FFT（Fast Fourier Transform）的Transformer模型，该模型可以处理长序列，并且在一些NLP任务上表现良好。
- Sparse Transformer：Sparse Transformer是一种使用稀疏注意力机制的Transformer模型，它可以减少Bert模型在处理长序列时的计算和存储成本。