

视听信息系统导论编程 3

高艺轩 毕嘉仪

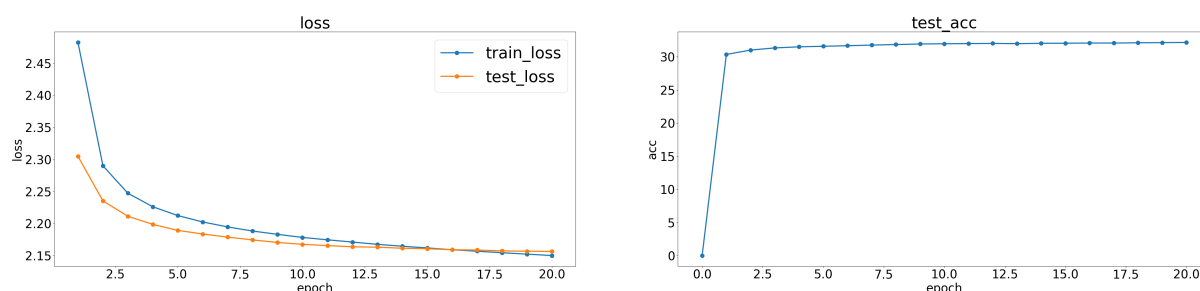
Task1 数据准备

函数`prepare_data`首先将全部文本读入，根据要求进行数据预处理，然后给每一个单独的字符都赋予一个 `index` 并将其存储为字典类型，得到对应词表。

Task2 完成 dataset

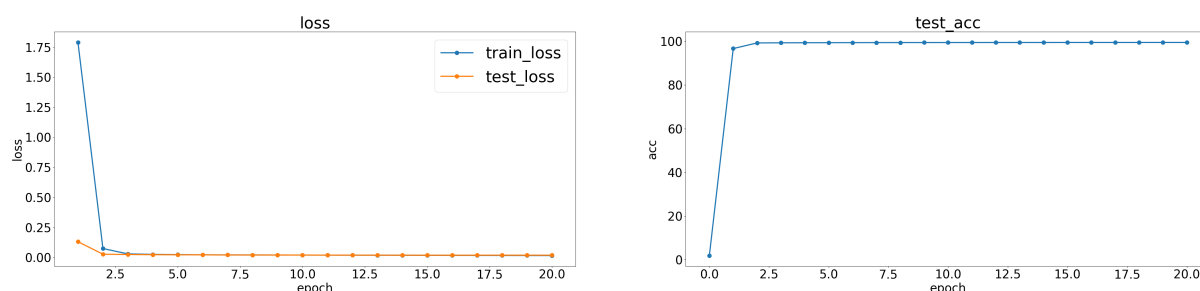
Task3 无 pos enc、无 attn mask 的 Transformer

训练结果：



Task4 有 pos enc、无 attn mask 的 Transformer

训练结果：

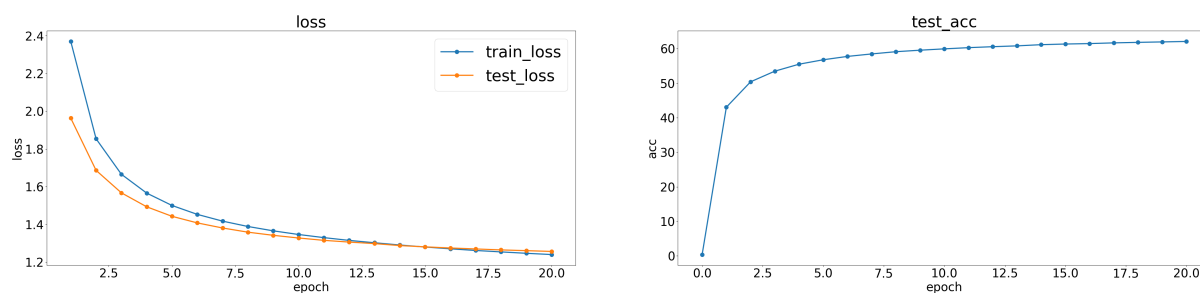


分析 positional encoding 的作用：

positional encoding 通过给每一个词向量添加一部分与位置有关的信息，能够让词向量携带单词在序列中所处位置的信息，从而提高模型对序列的理解能力。

Task5 有 pos enc、有 attn mask 的 Transformer

训练结果：



分析 attention mask 的作用：

attention mask 能够确定有效 token 的位置，即屏蔽掉词向量中不需要模型关注的部分（比如只看当前词语左侧序列，屏蔽右侧序列），从而提高模型的性能和准确性。

Task6 默认参数生成

Task7 不同 temperature 和不同 strategy 下生成内容的不同

程序源代码