样本生成方法、大语言模型微调方法及装置

说　明　书　摘　要

本说明书实施例提供了一种样本生成方法、大语言模型微调方法及装置。在该方法中，通过实时采集得到数据集，并对其中某个问题进行同义改写，将得到的多个同义问题分别构建成与下游任务对应的提示数据，将得到的多个提示数据分别输入大语言模型，得到多个预测结果。接着，基于多个预测结果确定大语言模型在下游任务中对第一问题文本进行预测时的确定性程度，利用该确定性程度确定多个提示数据的标注信息，进而可以利用标注信息和多个提示数据对所述大语言模型进行微调。采集得到的数据包含用户隐私数据，需要进行隐私保护。

图 2