

[发明公布] 语言模型训练方法、奖励模型训练方法、装置及电子设备

- 申请公布号：

CN117273117A

- 申请公布日：

2023.12.22

- 申请号：

2023112814194

- 申请日：

2023.09.28

- 申请人：

北京百度网讯科技有限公司

- 发明人：

柴业坤;王硕寰;孙宇

- 地址：

100085北京市海淀区上地十街10号百度大厦二层

- 分类号：

G06N3/092(2023.01)I; [全部](#)

- 摘要：

本公开提供了语言模型训练方法、奖励模型训练方法、装置及电子设备，涉及人工智能技术领域，具体为深度学习、自然语言处理、大模型等技术领域，可应用于基于人工智能的交互场景。具体实现方案为：获取语言模型基于目标问题所生成的回复；根据目标问题，调用目标工具，得到目标问题的参考信息，参考信息是基于目标工具的返回数据确定；根据目标问题的参考信息，对回复生成目标评价信息；根据目标评价信息，在强化学习阶段对语言模型进行模型参数调整。通过目标工具得到的参考信息中包含与目标问题的期望回复相关的信息，基于参考信息可以对回复进行准确客观地评价，因此，根据目标评价信息对语言模型进行模型参数调整可以提升语言模型的训练效果。[收起](#)



