

# 5 自定义微调

在本章中，我们将探讨自定义微调 AI 大语言模型的过程，微调是为了调整这些强大的 AI 语言模型以更好地满足您的特定需求，从而提高目标任务的性能。我们将以 GPT 或 ChatGPT 模型为例介绍微调模型的过程，主要包括准备自定义数据集、微调模型以及评估微调模型结果的基本步骤。

## 5.1 准备数据集

要微调 GPT 或 ChatGPT 模型，您首先需要一个代表目标任务的自定义数据集。理想情况下，该数据集应包含多样化、高质量的示例，以帮助模型了解您的特定应用程序的细微差别。准备自定义数据集的过程涉及几个关键步骤：

- (1) **数据收集：**收集反映任务性质的数据。这可以包括来自特定领域的文档、网站或现有数据库的文本。确保您有权使用这些数据，并且它符合道德考虑和隐私法规。
- (2) **数据预处理：**通过删除不相关或冗余的信息、更正错误和标准化格式来清理数据。此步骤对于确保您的数据集具有高质量且适合微调至关重要。
- (3) **数据注释：**通过标记示例、提供基本事实信息或指定所需输出来注释数据。此步骤可能需要专业知识或与领域专家合作以确保准确的注释。
- (4) **数据标注：**对数据集进行标注，为输入文本分配对应的目标输出。在监督学习任务中，这是关键的一步，因为模型将根据这些标签进行训练。
- (5) **数据增强（可选）：**通过对现有数据进行变换（如改变词序、使用同义词等）来扩展数据集，以提高模型的泛化能力。
- (6) **数据划分：**分配数据集的大部分数据（通常约 80%）用于训练模型，训练集将用于微调，分配数据集的一部分数据（通常约 10%）用于验证。验证集用于调整超参数和评估过拟合风险，分配数据集的剩余数据（通常约 10%）用于测试。测试集将在微调完成后用于评估模型在未知数据上的性能。

## 5.2 微调模型的步骤

准备好自定义数据集后，就可以继续微调 GPT 或 ChatGPT 模型。此过程涉及使用您的数据集调整模型的权重，以优化其在目标任务上的性能。按照以下步骤微调您的模型：

- (1) 选择预训练模型：根据您的需求选择合适的 GPT 或 ChatGPT 模型。在做出决定时，请考虑模型大小、推理速度和资源限制等因素。
- (2) 定义微调目标：确定您希望模型实现的具体目标，例如文本分类、摘要或问答。这将指导您的微调过程并帮助您建立适当的评估指标。
- (3) 配置超参数：为微调过程设置适当的学习率、批次大小和训练周期数。这些超参数会显著影响微调模型的性能，因此请考虑尝试不同的值并使用验证集来选择最佳配置。
- (4) 微调模型：利用自定义的数据集对模型的权重进行更新。在这个过程中要避免过拟合的问题，即模型对训练数据拟合过度，导致在新的、未见过的数据上性能表现不佳。可以通过正则化技术和早停策略来减轻过拟合的影响。

## 5. 3 评估微调模型

微调模型后，您需要评估其在目标任务上的性能。这涉及使用您在数据划分过程中保留的测试集。评估微调模型需要几个关键步骤：

- (1) 执行测试：使用经过微调的模型对测试集进行预测，生成相应的预测结果。此集合包含模型在训练或验证期间未见过的示例，使您可以评估其在新的、未见过的数据上的性能。
- (2) 选择评估指标：选择符合您的任务和微调目标的适当指标。根据目标任务的性质，选择合适的评估指标。例如，对于文本分类任务，可以使用准确率（accuracy）；对于摘要生成任务，可以使用 ROUGE 分数；对于翻译任务，可以使用 BLEU 分数。需要注意，没有任何一个指标可以完美地捕捉模型的性能，因此请考虑使用多个指标进行综合评估。
- (3) 计算评估分数：将模型的预测与测试集中的真实标签或输出进行比较。根据此比较计算选定的评估指标，以量化微调模型的性能。
- (4) 分析结果：检查模型在不同评估指标中的表现，并确定它擅长或困难的领域。对比评估指标的得分，分析模型在不同方面的性能。同时，关注模型在特定场景下的表现，检查模型是否存在偏见、过拟合等问题。
- (5) 迭代和优化：基于评估结果，确定模型性能存在的问题，并采取相应措施进行优化。这可能包括调整模型结构、优化训练参数、改进数据预处理方法等。然后重新微调模型，并在测试集上重新评估，直到达到满意的性能水平。

## 5. 4 微调模型示例

在本节中，将介绍有关如何为不同应用程序准备自定义数据集和微调 GPT 和 ChatGPT 模型的示例，包括情绪分析、摘要生成和特定领域的问答。

## 示例 1：情感分析

目标任务：对输入文本进行正面或负面情感分析。

准备自定义数据集：

- 收集原始数据：从在线平台收集电影或产品评论，确保数据集多样且足够大以覆盖各种情感表达。
- 标签数据：使用情绪标签手动注释评论，例如“正面”“负面”或“中性”。
- 格式化数据：将数据结构化为 JSON 或 CSV，包括评论文本和相关的情绪标签。
- 拆分数据：将数据分为训练集、验证集和测试集（例如，80% 用于训练，10% 用于验证，10% 用于测试）。

微调 GPT 或 ChatGPT 模型：

- 定义微调目标：训练模型根据评论文本预测情感标签。
- 选择超参数：选择合适的学习率、批量大小和训练周期数进行微调。
- 执行微调：使用训练集微调模型并在此过程中监控其在验证集上的性能。
- 评估模型：使用验证集评估微调后的模型性能。针对文本分类任务，可以采用准确率作为评估指标。根据结果，适时调整超参数，进一步优化模型。

## 示例 2：摘要生成

目标任务：从输入的长文本中生成简洁的摘要。

准备自定义数据集：收集一组长文本及其相应的摘要，如新闻文章、科研论文等。将数据分为训练集、验证集和测试集。确保数据集具有多样性，涵盖各种主题和文体。

数据预处理：对长文本和摘要进行清洗、分词和编码。为模型创建可理解的输入和目标输出格式。

微调模型：将训练集输入到 GPT 或 ChatGPT 中，调整模型权重以适应摘要生成任务。在微调过程中，注意调整学习率、批次大小和训练周期等超参数。

### 示例 3：技术支持

对于技术支持场景，假设您正在与一家希望使用 GPT 帮助客户解决常见问题的软件公司合作。首先，您需要创建一个包含相关问题和答案的数据集，并确保数据集具有多样性，涵盖各种客户需求和场景。

数据集可以分为两列：“问题”和“答案”。这是数据集外观的一个小示例：

问题，答案

“如何重置我的密码？”，“要重置密码，请按照以下步骤操作：1. 转到登录页面。2. 单击“忘记密码”。3. 输入您的电子邮件地址。4. 单击“提交”。5. 检查您的电子邮件以获取密码重置链接。6. 按照电子邮件中的说明创建新密码。”

“为什么软件没有回应？”，“如果软件没有回应，请尝试以下步骤：1. 关闭计算机上运行的所有不必要的应用程序。2. 重新启动软件。3. 将软件更新到最新版本。4. 检查您的计算机的系统要求以确保兼容性。如果问题仍然存在，请联系我们的支持团队以获得进一步帮助。”

请注意在此示例中使用了 CSV 格式。CSV 格式是逗号分隔值，是包含在双引号（“”）内的值以表示列值，同时以逗号（，）分隔以分隔列。不同的大语言模型可能支持不同的格式，需要注意根据不同的大语言模型适配不同的数据集格式。

在数据集上对模型进行微调后，您可以将模型部署到实际应用场景，如在线帮助中心、客户支持聊天机器人或者电话中心自动回答系统，您可以使用这样的提示来获得相关回应：

我在使用软件时遇到问题。它一直在崩溃。我该怎么做才能解决这个问题？

### 示例 4：食谱推荐

对于食谱推荐场景，您需要创建一个包含食材及其对应食谱的数据集。数据集的组织方式应使模型能够理解成分和食谱之间的关系。

#### 数据集

收集各种食谱，包括菜名、材料列表、制作步骤和烹饪时间等信息。这些数据可以从食谱网站、应用程序或专业书籍中获取。确保数据集具有丰富的食谱种类，涵盖各种口味、菜系和难度等级。这是一个小样本（你应该有数百、数千或更多）：

菜名：西红柿炖牛肉

材料：牛肉、西红柿、洋葱、大蒜、橄榄油、盐、胡椒

步骤：1. 将牛肉切成块，西红柿和洋葱切成丁，大蒜切末。2. 在锅中加入橄榄油，放入洋葱和大蒜翻炒至软。3. 加入牛肉块，炒至表面变色。4. 加入西红柿丁，继续翻炒。5. 加入适量水，盖上锅盖，炖煮 40 分钟。6. 加入盐和胡椒调味，煮 5 分钟即可。

烹饪时间：60 分钟

对模型进行微调后，您的 GPT 模型应该能够根据用户输入的关键字或喜好推荐相应的食谱。例如，当用户输入“素食”时，模型可以推荐适合素食者的食谱。如果用户输入“意大利面”，模型可以推荐各种意大利面食谱。

## 示例 5：医疗问答

对于医疗问答场景，需要一个包含医疗问题及其来自可靠来源的相应答案的数据集。数据集的组织方式应使模型能够理解问题和答案之间的关系。

### 数据集

这可能包括从权威医疗网站、医学教材、医学论文和专业医疗社区获取信息。请确保数据来源具有可靠性和准确性。将数据组织成输入（问题）和目标输出（答案）的形式。数据集可能包含多个问题答案对。这是一个小样本：

问题，答案

“普通感冒的症状是什么？”，“普通感冒症状包括：1. 流鼻涕或鼻塞。2. 喉咙痛。3. 咳嗽。4. 打喷嚏。5. 轻微头痛。6. 身体酸痛。7. 低烧。症状通常在接触引起感冒的病毒后一到三天出现，通常会持续 7 到 10 天。”

“感冒时可以吃抗生素吗？”，“感冒通常是由病毒引起的，抗生素只能治疗细菌感染，因此对感冒没有作用。在感冒时，应该多休息、多喝水，遵循医生的建议。”

在数据集上对模型进行微调后，您可以使用这样的提示来获得相关回应：

感冒有哪些症状？

这些示例演示了如何为各种应用程序准备自定义数据集和微调 GPT 和 ChatGPT 模型。请记住使用测试集彻底评估您的微调模型，以确保它们在看不见的数据上表现良好。此外，考虑重申微调过程并对超参数、数据集质量或模型架构进行调整，以进一步增强模型的性能。

通过执行这些步骤，可以创建更适合特定应用程序的自定义模型。微调本身是一个迭代过程，可能需要多次尝试才能获得最佳结果。可以考虑与领域专家合作，结合用户反馈，并及时了解技术最新进展。