**1.有很多個massage為什麼不合併成一個?**

格式應該跟Chat Completions API相同，其中不合併的原因是為每個messages都是一個獨立的對話示例，用於訓練模型如何在不同的對話情境下回應。例如:

{"messages": [{"role": "system", "content": "Marv is a factual chatbot that is also sarcastic."}, {"role": "user", "content": "What's the capital of France?"}, {"role": "assistant", "content": "Paris, as if everyone doesn't know that already."}]}

{"messages": [{"role": "system", "content": "Marv is a factual chatbot that is also sarcastic."}, {"role": "user", "content": "Who wrote 'Romeo and Juliet'?"}, {"role": "assistant", "content": "Oh, just some guy named William Shakespeare. Ever heard of him?"}]}

這兩段是不同的對話，如果將其合併在一起，會導致模型分不清楚上下文關係

**2.了解裡面Ft的格式要餵什麼?**

role:system content:(文本的風格或對話背景)

role:user content:(使用者的問題)

role: assistant content:(理想的回應)

weight:1 or 0(控制是否訓練該句(僅用在)assistant)確認沒寫是什麼(預設為1)

可以保留多久 多少

**3.怎麼讓他能夠講出對失智者的語氣**

先準備50個精心設計的對話(分為訓練和測試)

訓練完之後拿來跟其他模型對比，如果不如預期

 **針對剩餘問題收集示例**：

* 如果模型在某些方面仍然表現不佳，添加直接展示如何正確處理這些方面的訓練示例。

 **檢查現有示例是否有問題**：

* 如果模型存在語法、邏輯或風格問題，檢查數據中是否存在相同的問題。例如，如果模型錯誤地說「我會為你安排這個會議」，檢查是否有示例教模型說它可以做這些它實際不能做的事情。

 **考慮數據的平衡和多樣性**：

* 如果數據中60%的助手回應都是「我無法回答這個問題」，而在推理時只有5%的回應應該是這樣的，模型可能會過多地拒絕回答。

 **確保訓練示例包含所有必要的信息**：

* 如果希望模型根據用戶的個人特徵進行讚美，而訓練示例中助手的讚美是針對對話中未出現的特徵，模型可能會學會生成虛構的信息。

 **查看訓練示例的一致性**：

* 如果多個人創建了訓練數據，模型性能可能會受到人們之間一致性水平的限制。例如，在文本提取任務中，如果人們只對70%的提取片段達成一致，模型可能無法做得更好。

 **確保所有訓練示例格式一致**：

* 確保所有訓練示例的格式一致，與推理時期望的格式一致。

可以調整的參數:

* epochs（訓練次數）
* learning rate multiplier（學習率乘數）
* batch size（批次大小）(每次訓練步驟中使用的訓練樣本數量)

如果模型未按預期跟隨訓練數據，增加訓練次數1或2次。

如果模型變得不如預期的多樣性，減少訓練次數1或2次。

如果模型看起來沒有收斂，增加學習率乘數。

確認就是我要的，決定有幾種語氣，

**量表 或 引導運動 聊天 玩遊戲互動 吃飯 都先做在一起**

分三種情形

1.亂答的

2.不答的

3.驗證是否是對錯

(1)回答了但是錯的

(2)回答了確定是對的，接下來該如何引導下去

本週問題:

1.weight:1 or 0(控制是否訓練該句(僅用在)assistant)確認沒寫是什麼(預設為1)

2.句子有點難想，想到後面感覺質量也下降了很多，有沒有什麼地方可以取得大量且優質的對話資料(或者需要整組多一點人來想)(學習照顧失智症的人該怎麼講話互動)(短白話口語)(system要是keyword)\

(從A ft 一個 b，b的訓練資料要比較極端(白話簡短)，a有system b沒有進行比較，再把b也加上system三個一起比較,)

用短碼代表一句常用的話(user)

3.

* **量表**:我先挑了 MMSE認知功能評估量表裡的一些能用問答處理的問題來寫，剩下的需要現實中看護協助的我就還沒寫進去
* **引導運動:**有點不太清楚到時候的運動這塊該怎麼作，所以沒有寫很多
* **聊天:**這個部分是比較好寫的
* **遊戲:**遊戲都還沒做出來，我就還沒有寫這一塊
* **吃飯:**吃飯這邊到時候是要搭配每日紀錄的菜單嗎，那我該怎麼寫進ft裡

目標:gpt的性格、語調，或特定目的(跟rag有點交集)

4.

 **Training Loss**：0.8695

這是一個相對較低的損失值，表明模型在訓練數據上學習得很好。

 **Validation Loss**：2.0381

這個值明顯高於訓練損失，表明模型在驗證數據上的表現不如在訓練數據上的表現好。這可能意味著模型有些過度擬合?訓練數據。

 **Full Validation Loss**：1.3654

這個值介於訓練損失和驗證損失之間，更準確地反映了模型在整個驗證數據集上的性能。這表明模型在驗證數據集上的表現有所提升，但仍高於訓練損失。

5.

Ft模型

一張含有 文字, 螢幕擷取畫面, 字型, 軟體 的圖片

自動產生的描述

gpt3.5

一張含有 螢幕擷取畫面, 文字, 軟體, 多媒體軟體 的圖片

自動產生的描述

其餘部分都蠻糟糕…

1.有新的gpt 4o mini可以用，但要先付250美金，但微調的部分免費

字幕 聲音 畫面

1.先嘗試能否藉由聲音找出失智者的說話語音🡪對應字幕時間

2.找出失智者對答前後文(各問成2句)(注意去除重複)

3.標示互動應答的聲音/畫面/字幕起迄

m問n答

m=1,2,3

n=1,2

4.場景/話題分割