智慧課程問答系統

基於大語言模型的問答評論與輔助評分機制

摘要

本專題開發一個基於LINE聊天機器人的DS課程問答系統,旨在提升學生參與度與 教學管理的自動化。系統分為學生端與管理端,提供不同的操作體驗與功能服務。 系統透過Mistral模型修改prompt輸入,自動判斷學生答案的正確性,給予即時評分 與專業評語,並利用F1分數評估模型的精確性和穩定度。

特色:為使用者與管理者提供不同介面

使用者介面 開始作答=>選擇單元=>選擇題目=>收到評分與評論

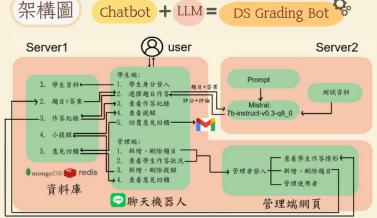


管理者介面(網頁)



語言模型:調整輸入 Prompt,採用 Mistral:7b-instruct-v0.3-q8_0 模型

由於評分及評論需求,並非單純的Classification或Evaluation, 因此我們在 Prompt 設計上參考了 Mistral AI官方網站中的多個範例,學習其中語法並綜合使用,修改出符合標準及格式的Prompt。 因 Mistral 模型的五種支援語言不包含中文,我們選用英文進行敘述,使模型在理解上更為精確。 針對評分和評論標準,我們定義了明確的準則,包括分數範圍及每個分數所對應的回答標準,藉此 減少模型在判斷時的混淆,提升評估的準確性與一致性。



評估



=> F1-score評估模型評分結果 將評分視為一個分類問題,利用混 淆矩陣計算出量化指標Precision和 Recall,從而再計算出F1-score。

混淆矩律		實際			
		0分	1分	2分	3分
預測	0分	預測正確	預測錯誤	預測錯誤	預測錯誤
	1分	預測錯誤	預測正確	預測錯誤	預測錯誤
	2分	預測錯誤	預測錯誤	預測正確	預測錯誤
	3分	預測錯誤	預測錯誤	預測錯誤	預測正確