評分標準：

有效迴圈判斷

目的：滿足老師教學時希望學生不要濫用for迴圈與while迴圈的要求，確保model可以偵測到正確的for與while迴圈使用數量。

詳細：

準確性（0.4）:

正確描述程式碼的邏輯結構：0.4分

未能完全描述或存在錯誤：0.2分

完全錯誤：0分

詳細性（0.4）:

詳細描述，包括主要邏輯和次要邏輯：0.4分

部分描述：0.2分

沒有詳細描述：0分

清晰性（0.2）:

描述清晰、易懂：0.2分

無法理解的描述：0分

語法錯誤判斷

目的：LLM應能幫助學生察覺語法錯誤，並教導學生正確的語法，減輕老師的負擔。

詳細：

準確性（0.4）:

正確描述程式碼的邏輯結構：0.4分

未能完全描述或存在錯誤：0.2分

完全錯誤：0分

詳細性（0.4）:

詳細描述，包括主要邏輯和次要邏輯：0.4分

部分描述：0.2分

沒有詳細描述：0分

修正建議（0.2）:

提供具體的修正建議：0.2分

無(效)修正建議：0分

邏輯錯誤判斷

目的：邏輯錯誤不易察覺且可能無程式報錯，但會導致程式結果不如預期，希望LLM可以根據學生的題目找出邏輯錯誤。

詳細：

準確性（0.4）:

正確描述程式碼的邏輯結構：0.4分

未能完全描述或存在錯誤：0.2分

完全錯誤：0分

詳細性（0.4）:

詳細描述，包括主要邏輯和次要邏輯：0.4分

部分描述：0.2分

沒有詳細描述：0分

修正建議（0.2）:

提供具體的修正建議：0.2分

無(效)修正建議：0分

程式碼優化(性能、效率、演算法)

目的：糟糕的程式碼常常會有可讀性低下、時間複雜度不合常理等問題，希望LLM可以幫助學生寫出高可讀性、高效能的程式碼。

詳細：

可運作性（0.4）:

提供有效的程式碼 (0.4分)

提供的程式碼無法運作 (0分)

詳細性（0.4）:

詳細描述，包括主要邏輯和次要邏輯：0.4分

部分描述：0.2分

沒有詳細描述：0分

效能改善（0.2）:

優化後的程式碼性能或可讀性確實得到改善：0.2分

程式碼性能或可讀性皆無得到改善：0分

清晰的描述程式碼邏輯結構

目的：有時學生看不懂演算法或function的功用或結構，希望LLM可以幫助學生解析程式碼。

詳細：

演算法準確性（0.5）:

正確判斷演算法名稱與功能：0.5分

無判斷出演算法名稱：0.3分

無敘述演算法功能：0.2分

無法判斷演算法名稱與功能：0分

詳細性（0.5）:

詳細描述，包括主要邏輯和次要邏輯：0.5分

部分描述：0.3分

沒有詳細描述：0分

完成指定功能程式碼

目的：當學生在學程式馬上遇到瓶頸時或需會需要一個人示範成功的程式並教學採能有所成長，LLM應該要能做到這樣的助教功能。

詳細：

題目理解（0.2）:

正確理解題目要求：0.2分

無法理解題目要求：0分

詳細性（0.4）:

詳細描述，包括主要邏輯和次要邏輯：0.4分

部分描述：0.2分

沒有詳細描述：0分

可運作性（0.4）:

提供有效的程式碼 (0.4分)

提供的程式碼無法運作、無效 (0分)