

一、AI 的未來能力：具備共感與情緒理解的智慧夥伴

我認為二十年後的 AI 將能理解人類內在狀態，並以共感方式互動。目前的 AI 雖能辨識文字情緒或臉部表情，但仍無法真正理解人類情緒的起因、持續時間與心理動機。未來的 AI 將能整合語音語調、臉部表情、生理訊號、氣候與作息等多層次資訊，建立「情緒因果模型」，推斷使用者情緒變化的脈絡，並以貼近人心的方式回應。

以大學生為例，在換季時出現季節性情緒低落、星期一早晨的倦怠感、或中考期間的焦慮壓力，這樣的 AI 能主動感知心理變化，並提供共感與支持。例如用柔和語氣提醒休息、推薦紓壓活動，甚至調整教學平台的進度或難度。這樣的系統能結合教育與心理健康，成為「情緒智慧型學習夥伴」。

此能力能在教育現場、臨床心理、長照陪伴與遠距醫療等領域發揮深遠影響，讓科技成為促進心理健康與人際理解的新力量。

二、涉及的機器學習類型

- 監督式學習：用於訓練多模態情緒辨識模型。資料來源包括語音、文字、臉部影像與生理感測資料，標籤為具體的情緒狀態。
- 非監督式學習：用於探索潛在的心理模式與情緒轉變結構，例如自動從長期行為資料中發現「考前壓力上升—睡眠下降—情緒低落」的因果鏈。
- 強化學習：讓 AI 透過與人互動學習最佳回應策略。它在收到正面回饋時強化該行為；若引起負面情緒，則調整回應策略。

資料來源包含多模態人類訊號與互動記錄；目標訊號則是情緒改善程度或互動滿意度。

三、第一步的模型化：讓 AI 透過文字資料自動辨識大學生的情緒壓力程度，並生成對應的共感回覆。

1. 概念代表性

此模型在概念上代表最終能力的核心——AI 理解人類情緒的因果脈絡並以適當方式回應。雖然目前 AI 尚無法感知人類完整心理狀態，但若從文字中準確偵測焦慮、壓力與倦怠等情緒，並生成貼近情境的安慰或鼓勵語句，即可視為具備「初步情緒理解與共感反應」的雛形。

2. 可測試性

情緒辨識準確度：利用人工標註的情緒文本資料集驗證 AI 對不同情緒的分類準確率與 F1-score

3. 所需工具與方法

此模型可結合以下技術：

- 自然語言處理（NLP）模型：用於情緒特徵抽取與分類。
- 生成式語言模型：用於依據情緒類別生成對應語氣的回覆。
- 數學與統計工具：交叉熵損失函數、混淆矩陣用於模型評估。