

比較分析報告

量化比較：

- 相似度將針對
 - a. "我今天心情很好，很適合去走一走"
 - b. "我最近覺得很開心"
 - c. "今天天氣讓我想要出去走一走"
 - d. "我打算加班到很晚"

做分析

- 摘要的原文為
"人工智慧（AI）的發展正深刻改變我們的生活方式。從早上起床時的智慧鬧鐘，到通勤時的路線規劃，再到工作中的各種輔助工具，AI無處不在。
在醫療領域，AI協助醫生進行疾病診斷，提高了診斷的準確率和效率。透過分析大量的醫療影像和病歷資料，AI能夠發現人眼容易忽略的細節，為患者提供更好的治療方案。
教育方面，AI個人化學習系統能夠根據每個學生的學習進度和特點，提供客製化的教學內容。這種因材施教的方式，讓學習變得更加高效和有趣。
然而，AI的快速發展也帶來了一些挑戰。首先是就業問題，許多傳統工作可能會被AI取代。其次是隱私和安全問題，AI系統需要大量數據來訓練，如何保護個人隱私成為重要議題。最後是倫理問題，AI的決策過程往往缺乏透明度，可能會產生偏見或歧視。
面對這些挑戰，我們需要在推動AI發展的同時，建立相應的法律法規和倫理準則。只有這樣，才能確保AI技術真正為人類福祉服務，創造一個更美好的未來。"
- 資訊保留度(評分重點)
(20%)概論：需提及人工智慧（AI）改變我們的生活方式
(20%)醫療：需提及提高了診斷的準確率和效率、能夠發現人眼容易忽略的細節
(20%)教育：需提及個人化學習系統、因材施教（客製化）準確率和效率
(40%)缺點：需提及就業、隱私和安全、倫理（策略過程往往缺乏透明度）
- 語句通順度(評分重點)
(20分)合理性(是否有自相矛盾)
(10分)有無冗詞贅字
(10分)連接詞使用是否通順

評估指標	TF-IDF(0-1)	GPT-4o(0-100)
相似度計算(ab/ac/bd/cd)	0.176/0.524/0.164/0.048	65/65/20/15
準確率	100%	100%
處理時間	0秒	2-3秒
成本	USD 0	USD 0.01
文本分類	可	可

支援類別數	有限	無限
自動摘要	須有關鍵字、統計高頻詞等等	可以直接輸prompt取摘要
資訊保留度	約70%	約85%
語句通順度	100分	100分
長度控制	困難	容易

質性分析

在分別做現代與傳統方法的文本比較，前者使用 open-ai 的 gpt-4o 模型分類與摘要，後者使用 TF-IDF 方法做文本相似度計算與透過給予分數與關鍵詞的方式分類與摘要。在實作過程中發現，傳統方法非常依賴原文的可分析程度，以「相似度計算」為例，當文本長度過小或設定的關鍵字出現頻率太低時，即使相似度較高的文本，其輸出的分數也會偏低；在摘要的部分也有相同的問題。但同時，傳統方法卻有更高的穩定性，程式也只針對設計者的設計產生輸出，反觀現代 AI 方法，較依靠使用者給予 prompt 的精確性與邏輯，太少的話，AI 會自行做延伸，太多又會多了很多不該出現的內容。不過，現代方法，的確節省許多時間，並不用花費篇幅撰寫「如何計算」的程式碼，可直接利用 prompt 進行指令的下達。我認為，現代 AI 方法應使用在文本的內容與背景主題等等較為不複雜的情況，而由於複雜的文本需要先天就設定好分類依據，因此適用傳統方法。

最為印象深刻的困難是在做傳統方法的「語意分析」，因為需要文本中詞對應到的情緒或主題完整列出，若太少就會導致分析不完全，太多將與輸入整篇文章沒什麼兩樣。但可以感受到傳統方法相當適用於「理性地」統計與分析，現代方法較適合「感性地」分析，後者比較容易考慮情緒。而在將來，會以「如何讓 AI 做理性地分析」為學習方向。

最後，傳統方法如 TF-IDF 適用於需快速處理、大量文本比對與關鍵字搜尋等場景，能以低成本提供可解釋的結果；而 AI 模型如 GPT-4o 則擅長語意理解、情緒分類與自動摘要，適合需深層語意分析與自然語言生成的應用。因此我們可先用 TF-IDF 預篩相關文本，再交由 AI 進行語意分析與摘要生成，兼具效率與準確性，達到低成本又高品質的文本處理與決策支援。

GitHub 連結

<https://github.com/Corngit/NTPU-AI-HW-2025Fall>