**一、摘要**

隨著電子商務的迅速發展，人們在對於查詢商品的需求也日益增加，當前主要的商品查詢方式為關鍵字與圖片搜索，但在某些時候，用戶對於產品的認識僅能以文字描述其樣貌以及功能，這在目前的常見查詢方式上沒辦法有效地找到用戶所描述的商品。

本研究以多商品組合搭配之推薦功能為目的，開發一套基於大語言模型（Large Language Model，簡稱為LLM）與RAG（Retrieval-Augmented Generation）的自然語言文本輸入之商品搜尋與推薦系統，以解決現有電子商務平台在組合式商品推薦上的不足。將使用者的自然語言商品描述需求透過BERT（Bidirectional Encoder Representations from Transformers）將輸入轉換至向量並進入向量資料庫中查詢，找出在商品資料庫中有符合需求的商品，並使用RAG模型進行該商品組合搭配以及文本生成，提供完整的商品搭配建議，滿足使用者在特定情境下的商品需求以及提升購物體驗。

關鍵詞：電子商務、多商品組合搭配、大語言模型、RAG、BERT

**二、研究動機與研究問題**

近年在電子商務以及科技的蓬勃發展下，商品種類的數量大幅增長，使消費者面臨的選擇困境日益加劇。哈佛研究證明人們有選擇障礙、不願意做出困難的決定，最大的原因在於害怕會後悔自己做的選擇[1]，而消費者也害怕自己會做出衝動性消費的行為，導致時常會發生在購買商品前到處詢問身旁好友以及網路上的商品評價，企圖說服這個選擇是跟隨大眾而不是自己的。而商品外型、性能、價格等需要在消費者心中具備獨特性，才能使消費者想購買此商品，在相同功能下，外型獨特更能討好消費者[2]。

目前普遍的商品搜尋系統大多依賴具體的關鍵詞或條件篩選，但當顧客無法明確提供商品名稱或類型時，使用此類系統往往變得較困難達成目的。在這種情況下，使用者可能需要在搜尋引擎（如 Google）上進行模糊查詢，或在社群軟體上發文求助。而大型語言模型（Large Language Model，簡稱為LLM）能夠解析使用者的模糊輸入，推測其需求並轉換為結構化的檢索條件，讓使用者不需要依賴精準的關鍵詞，也能夠達到其目的。

現有的研究或是計畫方案已能展示與聊天機器人互動進行模糊搜尋，如亞馬遜AWS部落格其中一篇文章，介紹如何使用亞馬遜的雲端服務建立智慧導購機器人，功能展示如圖1，能透過與機器人的聊天對話，幫助使用者找到符合情境的商品[3]，但即使成功找到相關商品，顧客仍然面臨從眾多選項中挑選最符合自身需求的難題。

在目前的研究展示中仍只能針對特定一項產品進行搜尋，無法應對組合式商品的推薦需求。例如，讀資訊工程系的學生想要組裝一台適合課業需求與經濟實惠的電腦，需考慮處理器、顯示卡、記憶體、硬碟等硬體的最佳搭配，或是希望穿著復古風的消費者，推薦系統不僅要考慮單件服飾，還需搭配褲子、鞋子、配件等，形成整體穿搭建議，又或是在居家室內裝潢時，需要在不同的空間下搭配沙發、桌椅、收納櫃等等家具以及滿足使用者的需求。這類涉及兩項以上商品的組合式推薦，以目前的系統在語意理解、需求解析及推薦邏輯上仍存在明顯的不足。

本研究以多商品組合搭配之推薦功能為目的，具備語意理解與模糊搜尋能力的AI商品搜尋系統，能根據顧客的自然語言描述轉換至文字向量庫中搜尋相關描述的商品，並能夠處理兩項以上的組合式商品推薦，根據使用者自身需求的情境下快速找到最適合的商品搭配組合，若使用者只想購買單項商品，也能夠提供額外資訊更好的搭配此商品，不只幫助使用者找到符合條件的商品，還能進一步的滿足使用者對於外觀的需求。

此研究除了能成為一種新型的商品搜尋功能之外，更是在未來儲存的資訊更加龐大且多元時，提供一種高效且智慧化的檢索方式，能運用至各種領域的搜尋系統中，縮短人們在「查找資料」所花費的時間，並進一步達到「資訊整合」的功能。

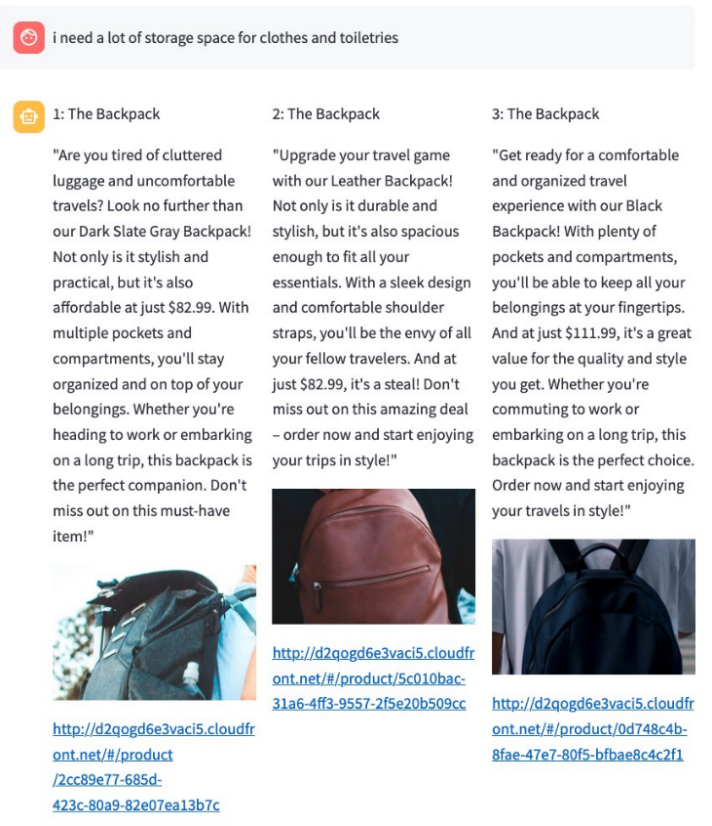


圖1. AWS部落格的智慧導購機器人方案Demo(來源[3])

**問題一：如何利用自然語言處理分析使用者需求**

透過結合LLM與分詞工具的前處理，系統能從用戶輸入的自然語言中提取關鍵特徵，並將其進行功能標註，並使用BERT理解文字中的語意與意圖，並可以從歷史對話來進一步調整搜尋結果，以便進行後續的檢索與推薦。

**問題二：多商品組合搭配功能如何幫助使用者?**

多商品組合搭配功能可以提升使用者的購物體驗，解決目前單一商品推薦的局限性。透過語意理解與使用者需求解析，系統能根據不同情境提供最佳的商品組合，減少使用者在搜尋與篩選上的負擔，在商品功能相同的情況下滿足使用者隱藏的商品外觀需求，幫助使用者更快速地做出決策，提升整體購物流程效率。

**三、文獻回顧與探討**

**3.1 向量資料庫（Vector Database）**

向量資料庫是一種專門用於存儲、檢索高維向量的資料庫，圖2以文字作為資料輸入表示，能夠將資料以數字矩陣向量的方式呈現，由於向量能夠以多維的方式表示，使其能夠將資料儲存於多維空間中並進行相似性搜索，適合處理非結構化數據（如文字、圖片、音訊等等）。

一張含有 文字, 圖表, 螢幕擷取畫面, 字型 的圖片

AI 產生的內容可能不正確。

圖2. 將輸入向量化後儲存至向量資料庫中[4]

**3.2 TF-IDF（Term Frequency–Inverse Document Frequency）**

TF-IDF是一種用於資訊檢索與文字探勘的常用加權技術，為一種統計方法，用來評估單詞對於文件的集合或詞庫中一份文件的重要程度。[5]

但在某些相似功能的詞是無法處理的，如「雨傘」「雨衣」是相似用途的商品，在此方法是根據詞語出現的頻率去計算相似度，導致兩詞語雖然意思相近，其相似性會是0，所以需要以商品的用途標籤進行此技術的處理，能夠優先篩選商品，在後續的檢索中提升效率。

**3.3 BERT模型**

BERT（Bidirectional Encoder Representations from Transformers）是由Google提出的一種基於Transformer架構的編碼器（encoder）[6][7]，能夠同時考慮單詞前後的語境，其訓練過程包含遮蔽語言模型（MLM）與下一句預測（NSP），能有效學習語意關聯。如圖3所示，BERT先在大規模語料庫進行預訓練，然後透過微調（Fine-Tuning）適應各種NLP任務，如問答、文本分類與語意匹配等，因廣泛運用多個NLP任務，使其衍生出多種變體，如RoBERTa和ALBERT，在開源平台Hugging Face's Transformers下能夠使用調整並應用其不同種類BERT模型。

一張含有 文字, 螢幕擷取畫面, 圖表, 字型 的圖片

AI 產生的內容可能不正確。

圖3. BERT的整體預訓練和微調

**3.4** **RAG（****Retrieval-Augmented Generation）檢索增強生成**

RAG是一種結合了搜尋檢索和生成能力的自然語言處理架構，模型可以從外部知識庫搜尋相關信息，然後在產生回應之前使用這些信息來生成回應或完成特定的NLP任務[8]。其流程架構圖如圖4：

(一)使用者輸入。

(二)將輸入傳送至檢索模組，從知識庫中查詢相關資訊。

(三)將檢索到的相關資訊傳回至系統形成增強語境（Enhanced Context）。

(四)原始查詢與增強後的資訊合併，再傳送至大型語言模型進行回應生成。

(五)根據合併後的資訊產生更完整且更有依據的回答。

一張含有 文字, 圖表, 螢幕擷取畫面, 方案 的圖片

AI 產生的內容可能不正確。

圖4. RAG與LLM的流程架構圖[8]

**四、研究方法及步驟**

本研究旨在開發一個自然語言處理且具備多項商品組合搭配之大語言模型推薦系統，如圖5架構所示，使用者輸入商品需求後，系統會透過前處理模組將輸入標準化，並從商品資料庫提取相關資訊，進一步由向量化模組將商品資訊轉換為向量。接著，檢索模組根據語意向量相似度查找符合類別的商品，並由篩選組合模組挑選適合的搭配組合。最後推薦解釋模組提供完整的商品資訊、推薦理由與相似需求問句，讓使用者能夠理解推薦內容。此研究將提供使用者更多商品搭配資訊與提升購物體驗。

**研究方法：**

一張含有 文字, 螢幕擷取畫面, 字型, 卡通 的圖片

AI 產生的內容可能不正確。

圖5. 研究流程架構圖

**研究架構步驟：**

**4.1 前處理模組**

從電子商務平台或商家提供的數據中收集商品種類、商品標題、描述、類別、價格、品牌等資訊，並去除停用詞、標點符號、數字和無關字元符號。使用分詞工具（如中研院-CKIP Lab）對中文文本進行分詞，將所有商品資訊整合成商品資料規範書，並儲存於關聯式資料庫中。同時也會透過微調（fine-tune）LLM進行商品自動化類別標註，例如，商品名稱為「GeForce RTX 4070 Ti SUPER」，LLM能夠自動標記為「電腦零件類」、「高效能」等標籤，幫助後續在選擇搭配時不會同時選擇相同功能及類型的商品進行組合。

在商品關聯性處理方面，系統亦可透過人工標註方式，為部分商品建立關聯標籤。例如，當商品為「筆記型電腦」時，可與「筆電包」、「滑鼠」、「外接螢幕」等相關商品建立關聯，又或是在家具類，將相同風格系列的商品建立專屬標籤，這些關聯資訊也會一併儲存在關聯式資料庫中，作為後續組合推薦的依據。此部分未來可透過4.4學習回饋模組進一步微調LLM增加自動化標註的占比。

**4.2 向量化模組**

將前處理後的商品資訊轉換為檢索模組可使用的向量表示，並儲存於向量資料庫中。對於前處理後的使用者輸入，系統會先應用TF-IDF方法來計算各單詞在整體商品資料庫中的重要性，提升關鍵詞的權重。而訓練過的BERT模型將商品資訊轉換為密集向量且根據剛剛計算出的關鍵詞進行加權調整，並透過語境化詞向量捕捉商品描述中的語意關係。舉例來說，透過TF-IDF計算出在輸入中的「電腦」、「衣服」、「椅子」等高相關性的商品關鍵詞，就會在BERT生成語意向量時透過向量元素相乘的方式進行加權調整，在後續的檢索模組中會更關注這些單字，最後則會產出查詢向量作為檢索模組的輸入。

**4.3 檢索模組**

查詢與商品向量的相似度將使用向量資料庫中的向量查詢功能（如FAISS），選出查詢與商品向量相似度相對接近的候選商品集，且將商品候選集中有「相同標註」的商品也一併列入提供給4.4章的篩選組合模組。例如，在資料庫中找到語意相對接近的商品是各種品牌的中階處理器、中高階顯示卡、電腦機殼等零組件，則會將這些零組件包含在關聯式資料庫中有相同標註的商品都取出，確保各式規格零件能夠被選出，達到多項商品組合搭配的先決條件。

**4.4 篩選組合模組**

系統會根據4.3節檢索出的商品，透過RAG技術檢查這些商品是否涵蓋了使用者需求所涉及的主要類別。例如，對於「讀資訊工程系的學生想要組裝一台適合課業需求與經濟實惠的電腦」輸入，商品組合應至少包含處理器、顯示卡、記憶體、主機板、電源供應器等核心零件。如果候選的商品類別不足，還是會保留此類別但輸出以沒有相關的產品做為告知，確保能提供給使用者完整商品組合需要的零組件資訊，而商品類別不足的情況將會記錄於4.6章回饋學習模組中，在未來進行模型優化避免同樣的問題發生。

後續篩選組合的步驟運用LLM的啟發式方法和RAG技術，自動生成不同的搭配方案，如「高性價比」、「均衡效能」、「高性能」等等方案，根據候選商品中的關鍵資訊與關聯標籤，調整並優化商品組合，確保同類型或相容性不匹配的商品不會出現於同一方案中，最終預計輸出3套完整的商品組合搭配方案。

**4.5 推薦解釋模組**

系統將透過LLM與RAG技術針對3套方案進行個人化商品推薦解釋與商品介紹文本的生成，讓使用者了解為何選擇這組商品進行搭配，也能在心中為這3套搭配進行自我評分，選出最適合自己需求的商品，就算只想購買一件商品的使用者也能知道如何透過其他商品的搭配更好的展示或使用。如圖6成品Demo所示，將輸出商品資訊、推薦理由文本以及可能的相似需求詢問來引導使用者是否還有其他需求未能解答?

**4.6 回饋學習模組**

(一)回饋數據收集機制

* 點擊行為：記錄使用者在推薦結果中點擊的商品，衡量哪些商品更符合需求
* 購買記錄：記錄使用者實際購買的商品，判斷哪些推薦更可能促成交易
* 手動調整：記錄使用者刪除或再次查詢的操作
* 商品評價：使用者對推薦商品的評價可作為強化模型的來源
* 商品類別不足的情況：在4.4章提到可能發生的篩選組合模組無法提供某類商品，系統將記錄並適時擴充商品資料庫與標註規則。

(二)模型更新技術

應用RLHF（Reinforcement Learning from Human Feedback）技術，使模型學習使用者的偏好。例如，當使用者調整推薦組合後，系統會根據「被選擇的商品」 與「被移除的商品」來調整推薦權重，並訓練策略模型來提升未來推薦的適應性

(三)持續強化商品資料標註與關聯性建模

現今商品種類的數量大幅增長，在4.4章中可能出現「商品類別不足」的情況，將透過 Zero-shot、One-shot、Few-shot的LLM標註技術，嘗試為新商品自動生成更完整的標籤。例如，當新款顯示卡上市但尚未被4.1章的模型標記為「高規格」，系統可自動從輸入的描述中推測其用途，並標記相關標籤，以改善未來的搭配推薦。

**4.7 前後台系統**

設計用戶前端界面，支持自然語言輸入與即時搜索建議，在獲得推薦結果後，輸出商品資訊、價格、搭配建議以及推薦理由。使用者亦可對推薦結果進行篩選、排序，或提供回饋，以進一步提升推薦品質。後台系統則負責處理向量化、檢索、推薦與回饋數據的管理，並提供商品管理功能，允許管理者手動標註商品關聯、調整標籤、更新商品資訊等。

**五、預期結果**

本系統將實現高效的商品搜索與組合商品推薦功能，提升用戶體驗與商家經營效率。用戶可透過自然語言描述快速找到最符合需求的商品，系統利用大語言模型進行語意理解，辨識同義詞、短語及上下文語境，提供準確的搜尋結果。對商家而言，使用者透過可互動式機器人查詢商品時能夠加以推廣自家品牌，增加商品曝光率，且也能夠成為線下店面銷售人員工具，使其更快理解客戶情境，並提供更多商品組合搭配的建議。

**六、需要指導教授指導內容**

本系統需要透過文本描述建立搜尋系統以及商品排名系統，也需要在資料探勘、資料庫設計、機器學習及推薦系統有豐富研究經驗的教授指導，需要在各階段驗證其功能完整性以及可用性等等，需要指導教授指導的部分如下內容：

* **機器學習模型開發**  
  系統需選擇適合的機器學習模型來處理商品文本數據，規劃完整的模型訓練流程，包括數據預處理、特徵提取與模型評估，並解決模型訓練中的過擬合問題，確保模型具備良好的泛化能力，以提升搜尋與排名的準確性。
* **推薦系統設計**  
  基於用戶行為數據（如點擊記錄、購買歷史）建立個性化推薦演算法，在準確性與多樣性之間取得平衡，以提升用戶滿意度。需要制定有效的評估方法，透過用戶回饋進行持續優化，確保推薦結果符合使用者需求並提升轉換率。
* **資料庫設計與優化**  
  設計高效的資料庫結構，以儲存商品資訊與向量化數據，優化查詢性能，以支援即時搜索與推薦功能，並且確保資料庫具備良好的擴展性與可維護性，以應對未來數據量的增長，確保系統長期穩定運行。

**七、參考文獻**

[1] **Barasz, K., Hagerty, S. F. (2021). Hoping for the worst? A paradoxical preference for bad news. Journal of Consumer Research, 48(2), 270-288.**

取自<https://doi.org/10.1093/jcr/ucab004>

[2] **MBA智庫百科-購買行為**

取自 <https://wiki.mbalib.com/zh-tw/%E8%B4%AD%E4%B9%B0%E8%A1%8C%E4%B8%BA>

[3] **AWS基於大語言模型與推薦系統建構電商智慧導購機器人**

取自 <https://aws.amazon.com/cn/blogs/china/build-an-e-commerce-intelligent-shopping-guide-robot-based-on-large-language-model-and-recommendation-system/?utm_source=chatgpt.com>

[4] **Best Open Source Vector Databases: A Comprehensive Guide**

取自 <https://www.graft.com/blog/top-open-source-vector-databases>

[5] **文字探勘之前處理與TF-IDF介紹**

取自 <https://www.cc.ntu.edu.tw/chinese/epaper/0031/20141220_3103.html>

[6] **Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.**

取自<https://arxiv.org/abs/1810.04805>

[7] **Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.**

取自 <https://arxiv.org/abs/1706.03762>

[8] **AWS-What is RAG (Retrieval-Augmented Generation)**

取自 <https://aws.amazon.com/what-is/retrieval-augmented-generation/>