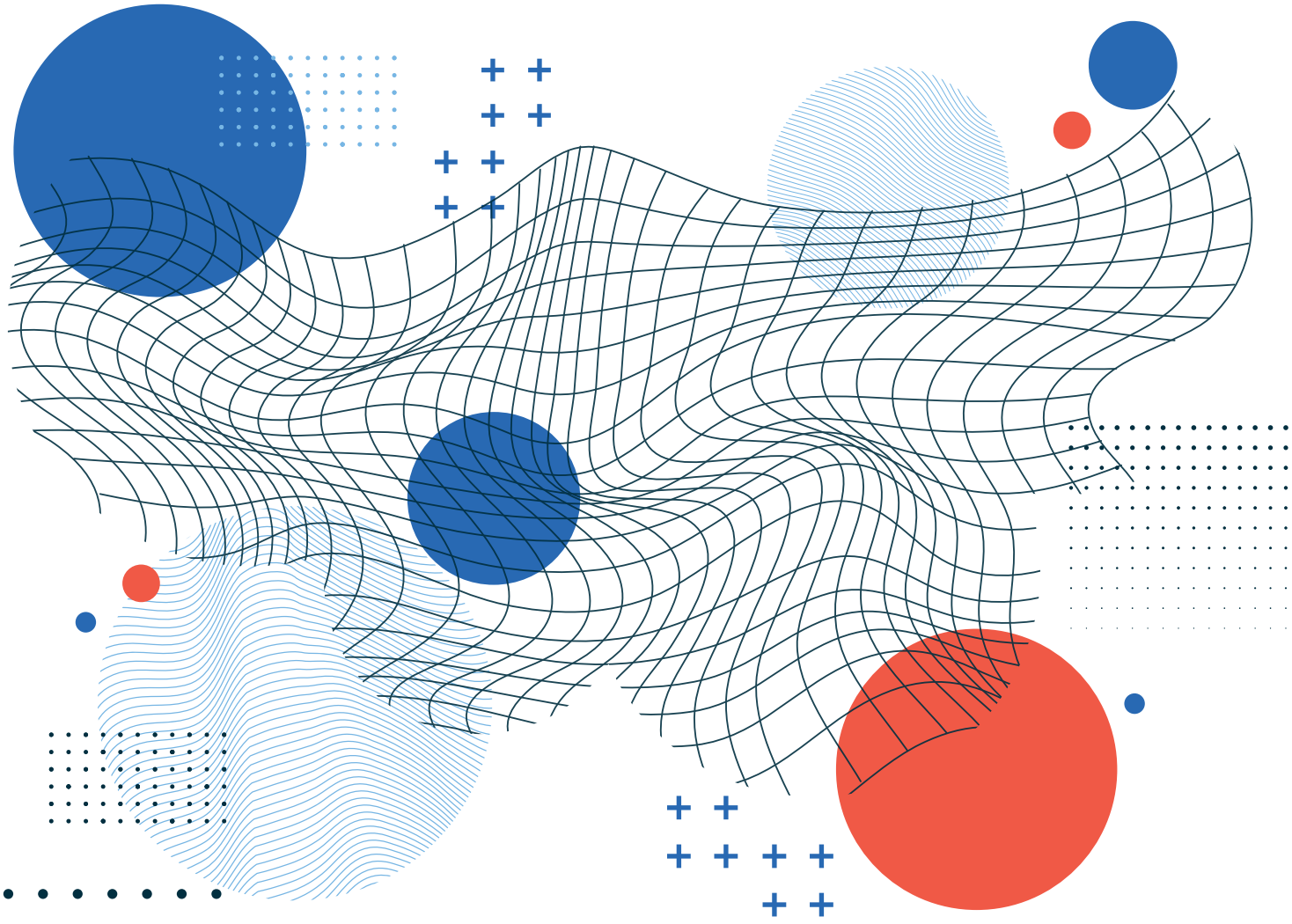
2026 ICEET

第十五屆 數位學習與教育科技國際研討會

International Conference on E-learning and Educational Technology

智慧教育×永續共榮

論文集 PROCEEDINGS



基於檢索增強雙重指令微調之《紅樓夢》古漢語知識問答系統建構

潘驄杰*
大學生
中原大學
資訊管理學系
E-mail：smartjay1206@gmail.com

賴錦慧
副教授
中原大學
資訊管理學系
E-mail：chlai@cycu.edu.tw

摘要

古漢語學習者常面臨閱讀文本時，背景知識不足、傳統檢索資訊方法效率較低、通用大語言模型易生幻覺與缺乏文獻出處之問題。基於此，本研究於單張 GPU 資源下，為古典文學名著《紅樓夢》建構古漢語知識問答之系統。為確保系統回答品質與精準度，使用檢索增強雙重指令微調架構（Retrieval-Augmented Dual Instruction Tuning, RA-DIT）微調 Qwen3-8B 模型以強化思考與拒答能力，最後評估教學輔助價值。實驗結果顯示：（1）檢索精準可靠：系統能有效融合關鍵字與語意線索，精準定位關鍵文獻段落，確保回答能溯源至原始文獻；（2）防幻覺機制穩健：落實「無文獻佐證即主動拒答」原則；（3）教學解析專業：生成之文學分析邏輯嚴密，其觀點與深度與專業評析契合，展現出實質古漢語知識輔助潛力。本研究證實於低計算資源條件下建構古漢語知識問答系統之可行性，並為後續研究者提供可復現之技術路徑。

關鍵字：

紅樓夢、檢索增強生成、檢索增強雙重指令微調、大型語言模型微調、古漢語（文言文）

＊為本文通訊作者

壹、緒論

古漢語（文言文）承載中華文化之精髓，為國文教育與人文研究之核心。然其學習門檻較高，學生常面臨三大困境：其一，因國學背景知識較匱乏，難以理解古漢語中之暗藏典故、歷史脈絡與哲思；其二，學習資源受限，針對查找專業內容，傳統檢索工具效率較低落（如紙本書籍），且師資輔導難以即時回應全體學生之個別需求；其三，雖通用大型語言模型（LLM）具強大生成潛力，卻因缺乏領域專屬知識與文獻溯源機制，易生「幻覺」（Hallucination）且無法提供可靠與方便之徵引文獻出處，難以滿足學術嚴謹性與教學信賴度。

鑑於上述問題背景，以及大語言模型訓練成本高昂，不利於一般教育機構或獨立研究者於有限算力下部署。為此，本研究以《紅樓夢》為典範場域，探索如何於單張 GPU 之低資源環境中，結合檢索增強生成（Retrieval-Augmented Generation, RAG）(Lewis et al., 2020) 與 RA-DIT 技術 (Lin et al., 2024)，建構一套低計算成本，且兼具文獻檢索、邏輯推理能力與防錯機制之古漢語知識問答系統，以方便學生使用高品質的學術問答系統，並為數位人文與 AI 輔助國文教學提供可復現之技術路徑。

因此，本研究確立三項核心目標：

- 導入檢索增強雙重指令微調（RA-DIT）架構，強化語言模型（Qwen3-8B）之邏輯推演，以及「無文獻證據即拒答」之防錯能力。
- 優化檢索增強生成（RAG）模組，結合混合檢索技術，實現在知識庫中，對複雜人物關係、典故、文脈之文獻段落精準定位與援引。
- 全面評估系統於事實正確性、引用忠實度與推理品質之表現，驗證其作為國文教學輔助工具之實務價值。

為達成上述研究目的，本研究採「知識庫建置、模型微調、檢索生成、多維評估」之實作研究流程。首先，輯錄《紅樓夢》全書、維基條目與相關國學典籍，經語意分塊與標籤化後建置向量知識庫。其次，以 Qwen3-8B 為基座模型，於 RA-DIT 架構下，進行語言模型微調 （Language Model Fine-Tuning, LM-ft），賦予其證據檢核與拒答能力。問答系統運作時，先於知識庫同步比對關鍵字與語意脈絡，精準定位相關文獻段落，經排序篩選後，再提供微調後之模型作為回答依據，確保生成內容皆有文獻支撐。

貳、文獻回顧

檢索增強生成（Retrieval-Augmented Generation, RAG）結合檢索系統與生成式模型，透過擷取外部知識庫文獻，輔助大型語言模型（LLMs）產出較精確且

具脈絡的回答，並減少因訓練資料有限而產生的「幻覺」，同時可隨資料庫更新而即時擴充知識 (Lewis et al., 2020)。

然而，RAG 中「檢索器找到的資料」不一定符合「生成模型真正需要的文獻」。RA-DIT 遂提出兩階段協同優化架構 (Lin et al., 2024)。第一階段為 LM-ft，系統將檢索片段當作背景提示，訓練模型學會從多篇資料中抓出關鍵證據，並忽略無關內容，必要時主動拒答，以降低幻覺風險。第二階段為檢索器微調（Retriever Fine-Tuning, R-ft），由已微調完成的語言模型擔任「評審」，對不同文獻片段給出「參考分數」（預測答案的機率），再以 KL 散度（Kullback-Leibler Divergence）衡量並縮小這些分數與檢索器原始排序之差距，使檢索器逐步學會語言模型偏好的證據類型，提升檢索與生成的整體默契。

在古漢語領域中，標註語料稀少但單篇文字往往蘊含大量典故與知識，呈現「低資源、富知識」的特性 （李紳、胡韌奮、王立軍，2024）。RA-DIT 透過上述「檢索與生成」的雙向對齊，能將龐大古籍轉化為可動態呼叫的知識來源，比從零訓練大型模型更具成本效益。

參、研究方法與設計

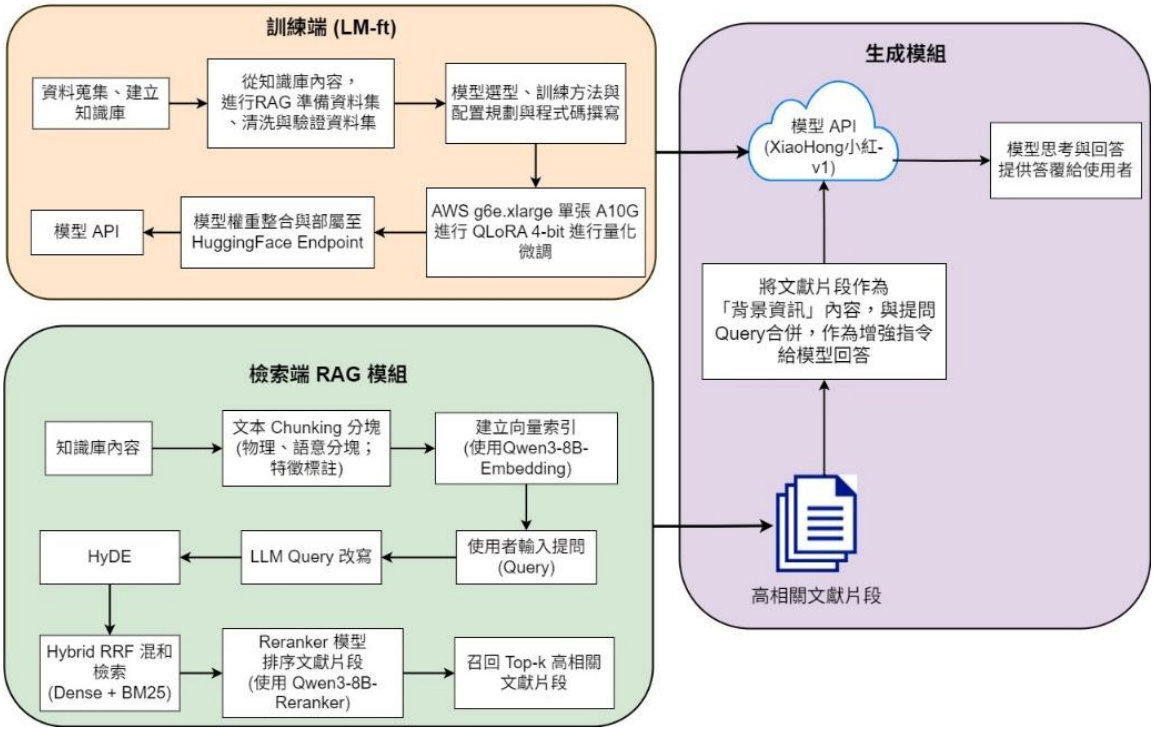


圖 1 紅樓夢之古漢語問答系統架構圖

一、訓練端：資料建構與模型優化（LM-ft）

如上頁圖 1 左上方區域所示，本模組涵蓋資料準備至模型微調之完整流程，核心目標在於約束基座模型建立嚴格依賴檢索文獻進行邏輯推演之運作機制；進而在低運算資源條件下，賦予模型深度學術解析能力，並建立防範幻覺之「無證據即拒答」防護機制。

- 外部知識庫建置：在檢索增強生成（RAG）外部知識庫之建置方面，本研究以確保文獻信度與文本溯源之精準度為核心要旨。系統完整收錄《紅樓夢》120 回原典文本與 102 首相關詩詞，更進一步整合 120 餘筆相關歷年學測與指考試題，並匯入 146 筆源自「經史子集」之結構化國學知識資料，以確立了知識來源之準確度與檢索範疇之完備性，強化系統針對深層文學隱喻之解析能力，並作為後續 RAG 填充背景文獻以準備訓練資料之基石。
- 訓練語料建立：基於 RA-DIT 架構之技術規範，本研究篩選並建立 4,332 筆基準黃金訓練資料集，進行指令微調 （instruction tuning）。該語料採用任務導向之問答對（QA pairs）格式，其結構高度集中於深度文學分析之訓練範本，混入約 20% 不相關文獻（負樣本）以建構主動拒答（Refusal）之對話情境，與嵌入思維鏈（Chain-of-Thought, CoT）推理標記，訓練模型輸出縝密之思考歷程。
- 模型微調與雲端部署：選用微調開源大型語言模型 Qwen3-8B，使用亞馬遜雲端運算服務（Amazon Web Services, AWS）之進行訓練。訓練後將模型命名為 XiaoHong-v1，並部署至 HuggingFace Endpoint，供後續模組即時呼叫使用 [1]。

二、檢索端：RAG 模組建置

如上頁圖 1 左下區域所示，本模組旨在於自建古漢語知識庫中檢索使用者相關問題之相關文檔。流程分為「離線索引建置」與「線上即時檢索」兩階段。考量完整檢索器微調（R-ft）之運算負載過重，且當代混合檢索與文獻重排技術佳，本研究基於單張 GPU 之資源限制，改以 RAG 模組機制替代，以兼顧檢索效能與低計算資源部署可行性。：

- 離線向量索引建置：該部分為建立知識庫，知識庫輯錄《紅樓夢》全書，及其維基條目與相關國學典籍。首先依文章結構與自然段落進行切割語意分塊（Semantic Chunking），並為各片段附加場景、人物等特徵標籤（Metadata）使其更依照文意切割。隨後透過嵌入模型（Embedding Model），其中使用 Qwen3-Embedding-8B 模型將分塊文本轉化為 3584 維之特徵向量（Embedding Vectors），而後建庫儲存以利後續之檢索使用。

[1] Hugging Face 模型專頁: https://huggingface.co/CongJ-Pan/XiaoHong-v1

- 線上即時檢索與重排：建立完知識庫後，當使用者輸入提問（Query），系統先執行提問改寫與擴充（Query Augmentation），並同步啟動假設性文檔嵌入（Hypothetical Document Embeddings, HyDE）機制，讓模型先以白話猜測可能的古漢語答案，再據此檢索，以彌補文言與白話文因表達方式差異所導致之檢索匹配困難。檢索階段同時啟動密集檢索（Dense Retrieval），捕捉整體語意氛圍，與稀疏檢索（BM25 Sparse Retrieval），鎖定專有名詞與精確詞彙，其初步結果經倒數排名融合（Reciprocal Rank Fusion, RRF）加權合併。最終交由重排序模型（Re-ranker Model），其中使用 Qwen3-8B-Reranker 模型進行精細交叉比對，篩選出最高相關度文獻（Top-k）遞交模型生成端。

三、生成端：防護機制與最終作答

如第三頁圖 1 右側區域所示，此階段負責整合檢索文獻與使用者提問，執行最終模型思考與正式輸出。

- 增強指令組裝：將檢索所得之高相關文獻片段作為參考資料，與原始提問拼接為增強提示詞（Augmented Prompt）。
- 模型思考與正式回答：增強提示詞傳送至 XiaoHong-v1 模型 API 後，進行思考流程，於此模型會優先評估文獻支撐度：若證據不足，則觸發訓練階段建立之防護機制，主動給出「拒絕回答」；若證據充分，則據文獻進行對比整合與邏輯思考後，產出結構嚴謹且具學術價值之正式回答。

最後，整體問答系統之原始碼已開源於 GitHub，且詳細記載其工程細節與系統介面實際呈現，方便後續研究者復現與體驗[2]。

四、系統介面實作展示

如下頁之圖 2，系統介面以深色為底，深紅為輔色之視覺化問答介面，營造專注沉浸的古典文學學習數位環境，該介面更提供使用者選擇「思考模式」與「RAG 模式」之切換樞紐，自主決定問答模式。再來，如圖 3 所示，當選擇 RAG 與思考模式後，系統顯示完整之「翻閱古籍索引」至「援引文獻」RAG 流程以及模型思考過程的推演步驟。而後，如圖 4 所示，系統於輸出回答解析後會同步陳列具體的文獻來源卡片，使用者能直接溯源實際由知識庫中的檢索片段內容，方便使用者做文獻回溯與查證。此外，更會依據當前對話脈絡自動生成「延伸推薦問題」，引導學生進行更深層次的探究學習。最後，如圖 5 所示，系統能主動攔截並婉拒數學運算或無意義字元等非紅學或國學相關提問，確保對話聚焦於古典文學主軸，此外，當面對知識庫未收錄之內容時，模型能透過思維鏈（CoT）明確觸發「無證據即拒答」之護欄機制，拒絕越界臆測與憑空捏造。

[2] GitHub 原始碼：https://github.com/CongJie-Pan/XiaoHong_RedChamber_QA_System

圖 2 系統主要入口介面

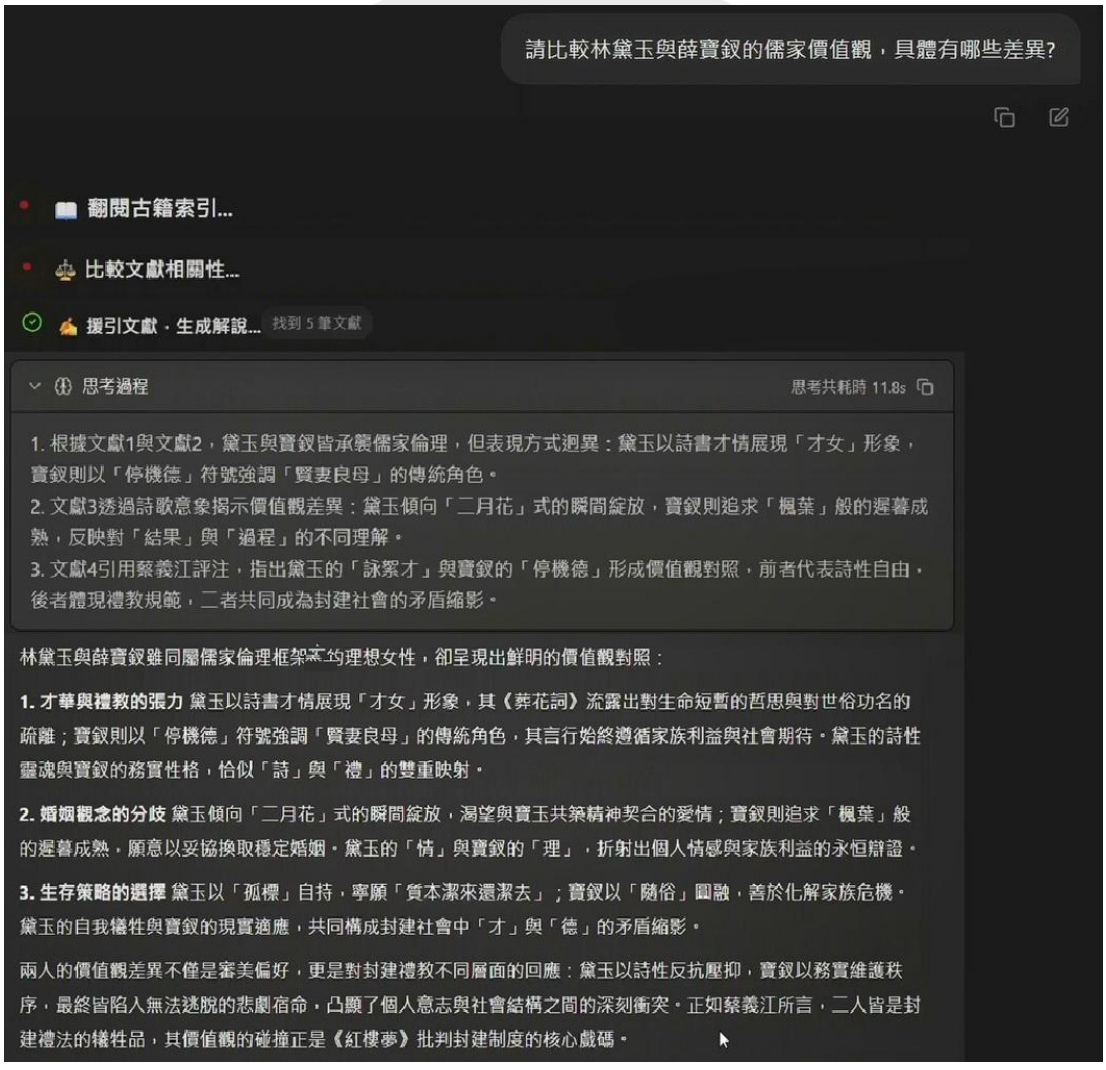


圖 3 系統進行 RAG 與思考模式後正式輸出回答之過程

圖 4 文獻來源卡片與延伸問題選擇

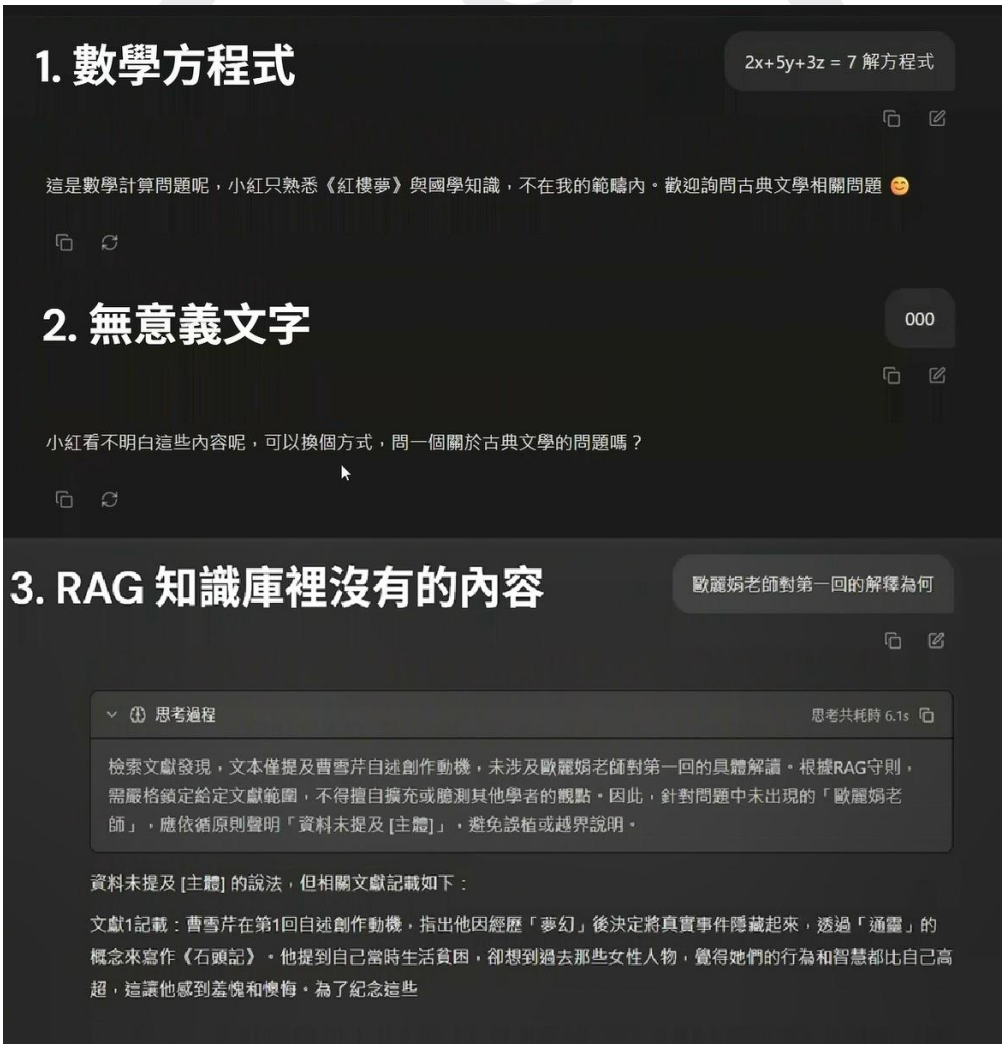


圖 5 系統拒絕回答之能力

肆、研究結果分析

本評估依布盧姆認知分類法（Bloom's Taxonomy） (Wilson & Leslie, 2016)，設計 60 題驗證集（事實題 40%、推理題 30%、綜合題 30%）評估系統整體之回答能力，公開於 HuggingFace 供學術復用 [3]。 本研究使用下列兩項指標：

- BERTScore (Zhang, Kishore, Wu, Weinberger, & Artzi, 2020)：基於預訓練語言模型將文本轉為向量，計算生成答案與參考答案之語意相似度(分數 0 – 1.0，越高表示語意越接近)，用以評估「答案核心意義是否正確」。
- 檢索增強生成評估框架 RAGAS (Es, James, Espinosa-Anke, & Schockaert, 2024)：透過比對生成內容與檢索文獻之對應關係，計算「忠實度」，答案是否可於文獻中找到依據，以及「脈絡召回率」，是否成功抓到關鍵文獻，以檢測「是否憑空捏造」。

評估架構設三組對照條件：「純模型」、「檢索增強（RAG）」、「檢索＋微調（RAG+LM-ft）」，旨在釐清外部文獻與模型邏輯訓練之各自貢獻。此外，為補足量化數據於文學解析之不足，導入 DeepSeek V3.2 與 Claude 4.6 Sonnet 語言模型擔任虛擬專家（LLM-as-a-Judge）(Zheng et al., 2023)，採雙盲主觀評分，由「事實正確性」、「引用忠實度」、「邏輯連貫性」、「教學語氣」等多維度進行主觀評量。

一、基礎模型與檢索增強之成效評估

如下頁圖 6 所示，實驗結果揭示了不同技術架構在特定題型上的效能取捨。首先，純模型下（Qwen3-8B）使用檢索增強技術（RAG）能有效約束模型的發散生成，顯著提升事實正確率並抑制幻覺（橘色 System B 事實題：0.7764 分）。而當模型進一步結合領域微調後配合 RAG（綠色 System C），由於微調訓練語料高度集中於深度文學分析與複雜推演邏輯，模型在權重更新的過程中產生了局部的「災難性遺忘」（Catastrophic Forgetting）現象，導致模型內部機率分佈發生偏移，使淺層事實記憶的提取能力微幅下滑（綠色 System C 事實題 0.7127 分）。

然而，此也使得複雜題型解析深度的大幅躍升（綠色 System C 綜合題 0.7549 分），質性案例分析亦印證此量化結果（完整對話比較詳見附錄一）：在「比較王熙鳳與賈探春」的綜合題中，相較於基礎模型常出現因果倒置與內容空泛的缺陷，System C 不僅能精確錨定原著中的關鍵情節，更能展現結合現代管理學視野與古典文學批評的跨學科深層剖析能力，精準對比兩人的治家手腕與性格差異。

[3] HuggingFace 驗證集：https://huggingface.co/datasets/CongJ-Pan/rcqa-system-XiaoHong-v1-golden-evalSets

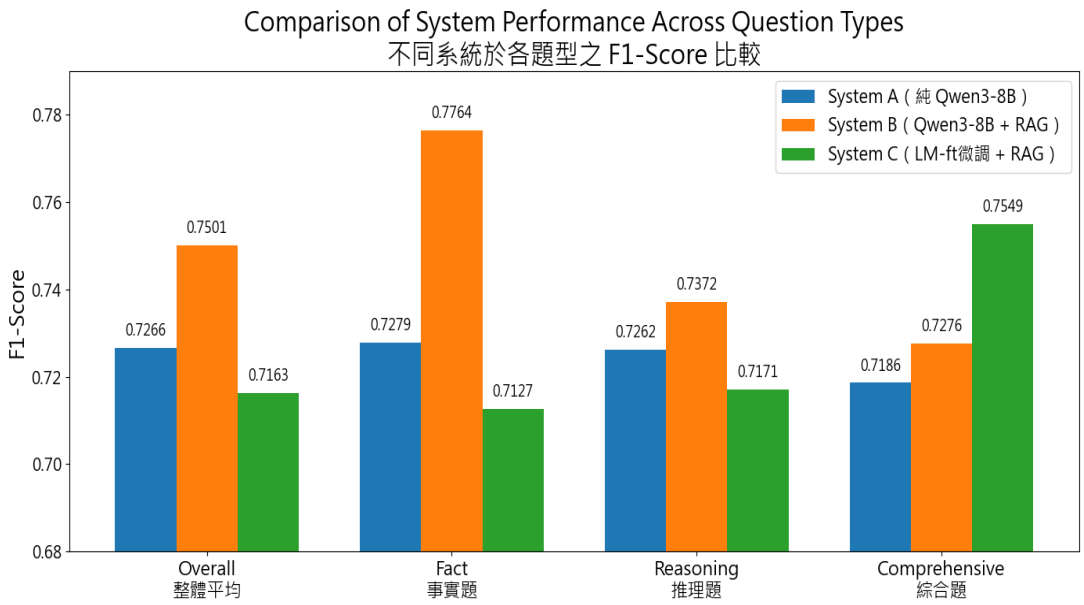


圖 6 純文本語意相似度（BERTScore F1-Score）之比較

二、檢索機制與參數設定之消融評估

如下表 1，透過逐一啟用或關閉不同檢索模組進行消融評估，結果證實引入外部檢索文獻能顯著提升答題準確率（開啟 RAG 搜尋：0.7141 分），且同時結合關鍵字與語意比對的混合策略最為有效（Hybrid RRF：0.7149 分）。針對古漢語文本特性，進一步發現外部文獻「精準少量」優於廣泛，僅保留最相關的一篇文獻即可有效過濾干擾雜訊，確保模型輸出緊扣文本脈絡。（僅取首：0.7254 分）

表 1 檢索流程與機制消融實驗之比較

實驗組別 (Ablation Target)	**測試配置(Configurations)**	**BERTScore F1-Score**
Exp A: 是否使用 RAG	關閉搜尋(純模型自己答)	0.6871
	開啟搜尋 (RAG 增強)	**0.7141**
Exp B: 搜尋方法	關鍵字 (BM25)	0.6916
	語意向量(Dense)	0.7148
	混合檢索(Hybrid RRF)	**0.7149**
Exp C: 給系統看幾篇最好的文獻 (Top-K 數量)	僅取首篇(Top-K=1)	**0.7254**
	取前三篇(Top-K=3)	0.7142
	取前五篇 (Top-K=5)	0.7141

三、生成語意相似度與文獻忠實度評估

如下圖 7，透過語意比對與文獻溯源雙重驗證，結果顯示系統生成之解析不僅核心觀點與專業評析高度吻合（平均召回率 BERTScore：0.9144），且在防幻覺機制上，能精準鎖定關鍵典籍段落，其生成內容皆具備明確出處（RAGAS 脈絡召回率：0.8174）、忠實度 0.6392（滿分 1.0） 則代表約六成句子可明確對應至原文，其餘四成未完全吻合之處，多屬模型對詩詞意境或哲理之合理推論與學術性延伸，非無憑據之幻覺。

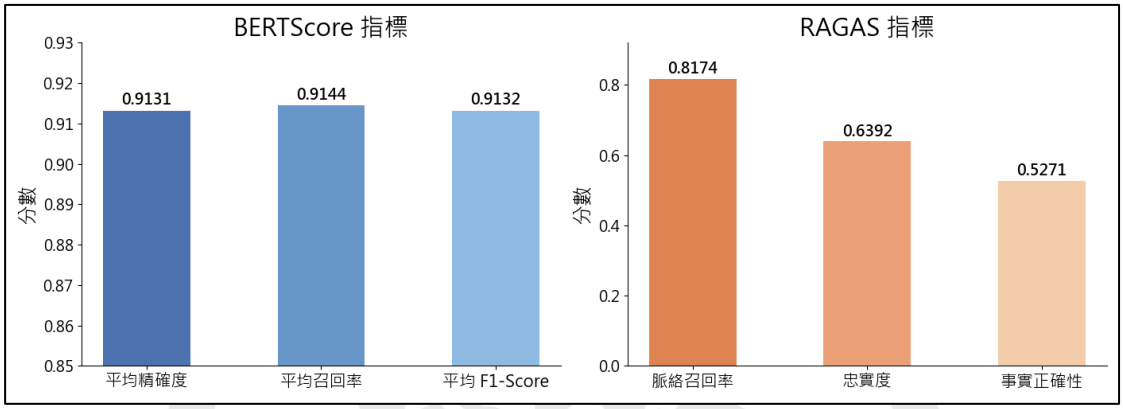


圖 7 BERTScore 語意相似度與 RAGAS 文獻忠實度指標之比較

四、教學輔助價值之主觀效能評估

如下頁之圖 8 ，本研究以中文語料為訓練核心之 DeepSeek V3.2，以及外文語料佔比較高之 Claude 4.6 Sonnet 作為虛擬裁判，採 1–5 分李克特量表進行五維度之雙盲交叉評估。在「事實正確性」上，系統常因《紅樓夢》龐雜之實體關聯（如寶玉與寶釵之人物信物錯配），觸發了模型嚴格的事實核查機制而遭扣分；然而在「忠實度」方面，雙方皆給予 4.1 分以上之高評價，顯示系統在微觀實體雖有偶發性誤差，但其生成內容仍高度依附於所檢索之文獻，有效抑制了嚴重幻覺的產生。

在「完整性」指標上兩者呈現顯著差異：具備較佳中文語境理解力的 DeepSeek V3.2（4.53 分），在面對文學深度分析時，即便生成文本未完全貼合制式的評分框架，仍能精準捕捉古漢語中的「言外之意」與深層隱喻，給予高度肯定；相對而言，Claude 4.6 Sonnet（3.67 分）則傾向採用剛性且線性的邏輯標準進行審核，導致評分偏低。同時，雙方在「教育實用性」與「邏輯推理」皆達成共識，肯定系統在解析複雜文學問題時所展現的推演邏輯，證明其具備導入課堂作為輔助教學工具之潛力。

總體而言，DeepSeek V3.2 的平均評分優於 Claude 4.6 Sonnet，此評分差異側面反映出：在評估古漢語與紅學等具備深厚文化底蘊之領域時，中文原生語言模型之評估機制與知識表徵，確實更契合該領域之文本特性。

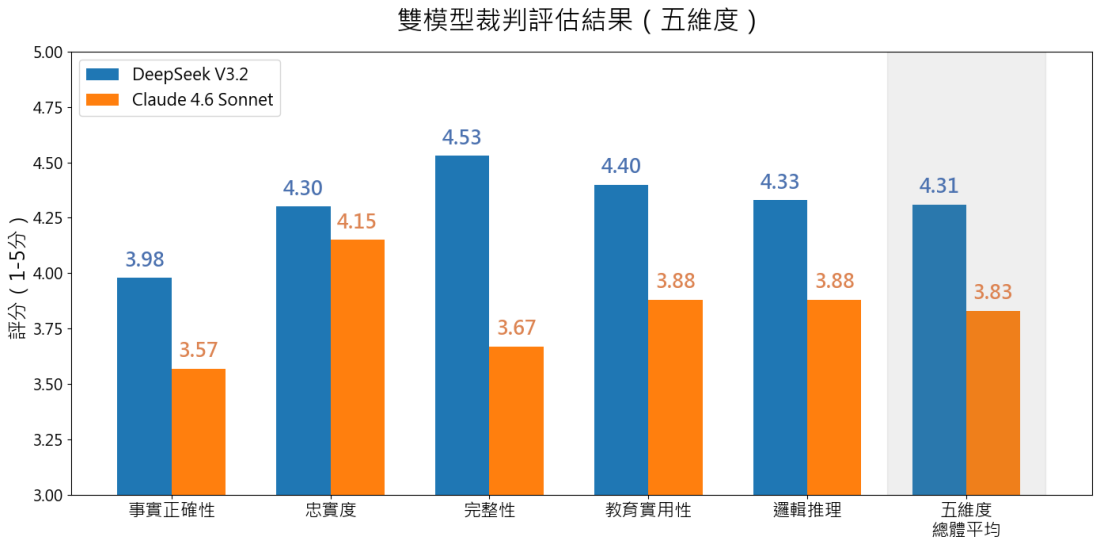


圖 8 兩大模型主觀評審結果（滿分 5.0）

伍、結論與未來研究

本研究旨在解決古漢語學習者面臨的背景知識匱乏、傳統檢索效率低落，以及通用大語言模型易生幻覺且缺乏出處等挑戰。受限於單張 GPU 之計算資源，本研究基於檢索增強雙重指令微調（RA-DIT）架構進行優化，成功建構具備文獻溯源與主動拒答能力之《紅樓夢》古漢語知識問答系統「小紅」。本系統之方法核心由兩大模組協同運作：檢索端（RAG）結合混合檢索策略，能於龐雜古籍知識庫中精準定位關鍵段落；訓練端（LM-ft）則透過指令微調強化模型之邏輯推演能力，並確立「無文獻佐證即主動拒答」之防護機制。

實驗結果中，本系統於三項核心評估面向均達預期目標：於檢索效能，混合檢索策略使關鍵文獻召回率達 81.74%，確保系統能精準找到逾八成所需文獻之段落；於防幻覺機制，RAGAS 忠實度達 0.6392 且 AI 虛擬評審忠實度平均分逾 4.1（滿分 5.0），驗證「無證據即拒答」之設計有效抑制憑空捏造；於教學應用，AI 虛擬評審於教育實用性評分，獲 4.14 分，顯示其解析兼具學術嚴謹性與課堂輔助潛力。整體而言，本研究驗證了低計算資源環境下建構高可靠度古漢語知識問答工具之可行性。

後續研究預計將朝 RAG 檢索優化、模型優化與實務驗證三面向進行。在 RAG 檢索優化上，未來擬繼續提升 RAG 文獻篩選的精準度與穩定度，提供穩定、精確之檢索結果；模型優化上，在指令微調階段強化「無佐證即拒答」之語料，以增強防幻覺能力，並加入引導式提問等多樣化教學情境語料，提升系統適應不同與學生使用者互動之能力；在實務驗證上，後續將邀請國文專家與師生實際試用，並導入高中課堂進行前、後測對照，評估其作為輔助教學工具之實際效益。

參考文獻

一、中文部分

李紳、胡韌奮、王立軍（2024）。古漢語大語言模型的構建及應用研究。**語言戰略研究**，*9* (5)，22-33。doi:10.19689/j.cnki.cn10-1361/h.20240502

二、英文部分

Es, S., James, J., Espinosa-Anke, L., & Schockaert, S. (2024). Ragas: Automated evaluation of retrieval augmented generation. *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., . . . Rocktäschel, T. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.

Lin, X. V., Chen, X., Chen, M., Shi, W., Lomeli, M., James, R., . . . Lewis, M. (2024). RA-DIT: Retrieval-augmented dual instruction tuning. arXiv:2310.01352.

Wilson, L. O., & Leslie, C. (2016). Anderson and Krathwohl Bloom's taxonomy revised: Understanding the new version of Bloom's taxonomy. *The Second Principle*.

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2020). BERTScore: Evaluating text generation with BERT. arXiv:1904.09675.

Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., . . . Stoica, I. (2023). Judging LLM-as-a-judge with MT-bench and chatbot arena. *Advances in Neural Information Processing Systems*, *36*, 46595-46623.

A Classical Chinese Knowledge QA System for Dream of the Red Chamber Using Retrieval-Augmented Dual Instruction Tuning

Cong-Jie Pan*
Undergraduate Student
Department of Information Management
Chung Yuan Christian University
Taoyuan City, Taiwan
E-mail: smartjay1206@gmail.com

Chin-Hui Lai
Associate Professor
Department of Information Management
Chung Yuan Christian University
Taoyuan City, Taiwan
E-mail: chlai@cycu.edu.tw

Abstract

Classical Chinese learners often struggle with limited background knowledge, inefficient retrieval, and LLM hallucinations. To address this, we developed a knowledge QA system for Dream of the Red Chamber under single-GPU constraints. By applying the Retrieval-Augmented Dual Instruction Tuning (RA-DIT) framework to fine-tune the Qwen3-8B model, we enhanced its reasoning and active refusal capabilities. Results demonstrate: (1) precise, fully traceable retrieval using combined keyword and semantic cues; (2) robust anti-hallucination mechanisms that actively refuse unverified queries; (3) rigorous pedagogical interpretations aligning with expert critiques. This study proves the feasibility of building accurate, low-compute educational QA systems and provides a reproducible technical pathway.

Keywords: Dream of the Red Chamber, Retrieval-Augmented Generation (RAG), Retrieval-Augmented Dual Instruction Tuning (RA-DIT), Large Language Model Fine-Tuning, Classical Chinese (Literary Chinese)

＊Corresponding Author

附錄

附錄一：微調前後對話實例比較

Query	王熙鳳與賈探春皆是《紅樓夢》中具有治事才幹的女性，且都曾在賈府的管理中扮演關鍵角色。請綜合比較兩人在治理賈府時的『地位基礎』、『治理風格與效果』，以及她們的『性別困境與歷史角色』有何根本差異？
模型微調前： System A (純 Qwen3-8B)	1. **地位基礎**：王熙鳳憑藉賈母信任與夫權（賈璉）掌握權力，屬「權力繼承」；探春則因才能與威望被賈府臨時委任，屬「才能突顯」。 2. **治理風格與效果**：王熙鳳以權術壓制、重利潤分配，短期維持秩序但加深腐敗；探春推行制度化改革，提升效率卻無法根治財政問題。 3. **性別困境與歷史角色**：王熙鳳雖掌權卻受制於夫權與家族規範，探春雖具才能仍受限於庶出身份與性別，二者皆無法突破封建體制對女性的束縛。 4. **衰敗的不同面向**：王熙鳳的權術反映賈府權力結構的腐敗，探春的改革失敗則揭示封建體制對女性才能的壓抑，二者共同說明賈府無法因內部治理而逆轉衰亡。
模型微調後： System C (LM-ft 微調+RAG)	### 一、地位基礎：血緣與制度的雙重差異 - **王熙鳳**：憑藉「內宅掌權」與「家族血緣」確立權威，她是榮府實際掌控者，擁有「鳳辣子」的稱號，權力來自於「嫡系長房」與「管家職責」的結合。 - **賈探春**：雖為庶出，卻因「王夫人信任」與「代理管家」的臨時授權，得以參與決策。她的權力基礎建立在「制度性代理」而非血統優勢之上，形成「非正式權力」與「正式權力」的錯位。 ### 二、治理風格與效果：威壓 vs 理性規劃 - **王熙鳳**：採用「威壓式」治理，強調「鐵腕手段」維持秩序，如「抄檢大觀園」時毫不留情，甚至不惜犧牲下人。她的治理效果短期有效，卻埋下「權力失衡」的長期風險。 - **賈探春**：展現「理性規劃」，推動大觀園改革，引入「承包制」與「獎罰分明」，試圖以制度設計提升效率。她的治理風格更接近現代管理，但受限於「臨時代理」與「家族政治」，難以徹底改變賈府頹勢。

	### 三、性別困境與歷史角色：壓抑 vs 傳承 - **王熙鳳**：作為「女性掌權者」，她被迫在男性權力結構中生存，最終走向「權力反噬」。她的「陰狠經濟學」暴露封建倫理的裂痕，卻也成為封建社會末期女性才能的縮影。 - **賈探春**：雖具備「治事長才」，卻因「庶出」與「女性」身份被排除在繼承序列之外。她的「遠嫁和親」象徵封建社會對女性才能的壓抑，同時也暗示「才能」與「命運」之間的不可調和。 ### 四、總結：制度與人性的交鋒 王熙鳳與賈探春的差異不僅在於個人能力，更體現封建社會中「血緣制度」與「女性角色」的雙重束縛。前者以「權力」為劍，後者以「智慧」為盾，共同折射出《紅樓夢》對封建末世的深刻批判。探春的「遠嫁」與鳳姐的「死亡」，皆是封建社會對女性才能的終極否定，卻也成為後世對「才能與命運」辯證的永恒啟示。

指導單位：國家科學及技術委員會
主辦單位：國立臺北教育大學資訊科學系｜中華資訊與科技教育學會 **承辦單位**：國立臺北教育大學資訊科學系｜中華資訊與科技教育學會
協辦單位：國立臺北教育大學人工智慧產學發展中心｜國立政治大學資訊科學系｜國立臺北教育大學課程與教學傳播科技研究所
台灣教育傳播暨科技學會｜碩陽數位科技有限公司
國家科學及技術委員會補助