
第六組 專題研究

探索 RAG 中關鍵字匹配與向量語
意搜尋結合的最佳權重配置

目次

03 分工表

04 RAG

06 研究動機

08 參考文獻

09 研究方法

10 資料說明

11 程式碼 + 研究結果展示

13 結論

15 未來展望

分工表

許新翎：RAG & 自動測試程式設計

黃威昊：測試資料製作

張元軍：測試資料製作

黃中璟：資料處理與分析

李逸誠：資料可視化與圖表

駱辰瑞：報告統整與簡報製作

王宏毅：報告統整與簡報製作



RAG (Retrieval-Augmented Generation)

基本結構：一個檢索器（retriever）和一個生成器（generator）

- 比喻：律師與助理
- 圖書館：向量資料庫
- 助理：檢索器（retriever）
- 律師：LLM（大型語言模型）
- * RAG + Finetune 的優勢 --> 上一學期的刑法課

Prompt (提示詞) 比較

使用者問題：

請幫我搜尋109學年吳文傑老師開設的全英授課經濟學課程

請幫我搜尋109學年吳文傑老師開設的全英授課經濟學課程



根據我的知識庫，我無法直接搜尋特定學年的課程資訊。如果你需要關於某位教授開設課程的詳細資訊，建議你查看學校的官方網站或課程手冊，或者直接聯繫學校的相關部門。這樣你可以得到最新和最準確的資料。

如果你有其他問題或需要幫助，隨時告訴我！



請從下列兩個選項中選擇出最適合回答 "我的問題" 的答案

"我的問題": 請幫我搜尋109學年吳文傑老師開設的全英授課經濟學課程

"選項一": title:111-1國立政治大學吳文傑經濟學,description:英語授課且是學年課，上學期是總經範圍下學期是個經範圍。主要會透過每週大量練習和詳細講解來建立經濟學概念，助教課時段很多老師也很多，適合透過土法煉鋼來奠定基礎的同學。 ,evaluation:好多好多英文,每週都有作業,老師很有魅力

"選項二": title:109-1國立政治大學吳文傑經濟學,description:全英文授課，政大少數全英文教學還能教得那麼好的老師，課程有很多加分的機會，老師會用小組討論的方式讓同學們討論，想練習英文或修ETP學程的同學極推，非商學院也可以來修，對經濟學能扎實的打下基礎。期末有個分組英文拍片的報告，也很有趣，不用怕會講不好或給分很低，老師人很好，大會考也都基本上是課程有教到的內容。 ,evaluation:人超多課很難選,Jack人超Nice,全英文

輸出: 請將選中的那個選項內的完整課程名稱,內容,課程介紹等 *所有資訊* 都輸出出來



"選項二": title:109-1國立政治大學吳文傑經濟學,description:全英文授課，政大少數全英文教學還能教得那麼好的老師，課程有很多加分的機會，老師會用小組討論的方式讓同學們討論，想練習英文或修ETP學程的同學極推，非商學院也可以來修，對經濟學能扎實的打下基礎。期末有個分組英文拍片的報告，也很有趣，不用怕會講不好或給分很低，老師人很好，大會考也都基本上是課程有教到的內容。 ,evaluation:人超多課很難選,Jack人超Nice,全英文

研究動機

RAG 準確率探究

2022 後 GenAI 新趨勢

傳統 AI 研究不容易想到突破

對實際產業界與學界的重要性



提出問題

Q：RAG 只靠語意檢索不行？ A：Hybrid Search

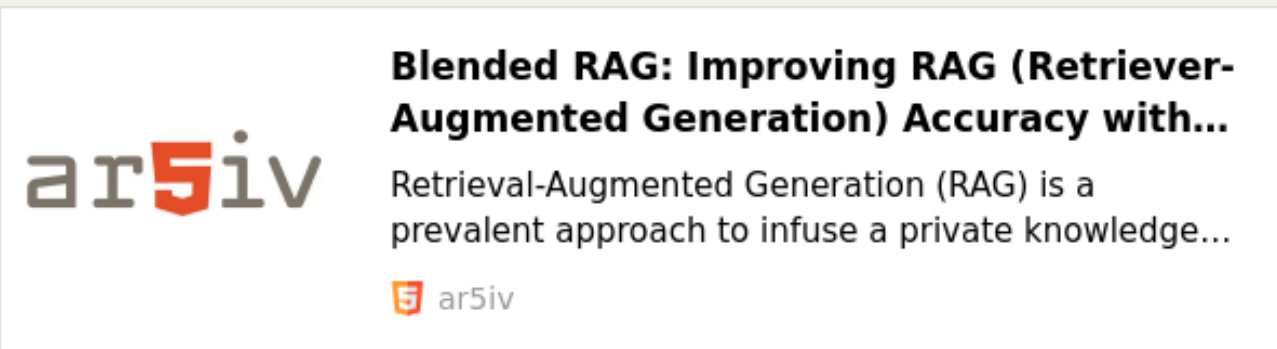
- 藉由 微軟與 IBM 論文 得到的啟發 -->
尋找另一種檢索合作：關鍵字檢索 (BM25F)

選擇：Weaviate 調整 Alpha 做 Hybrid Search

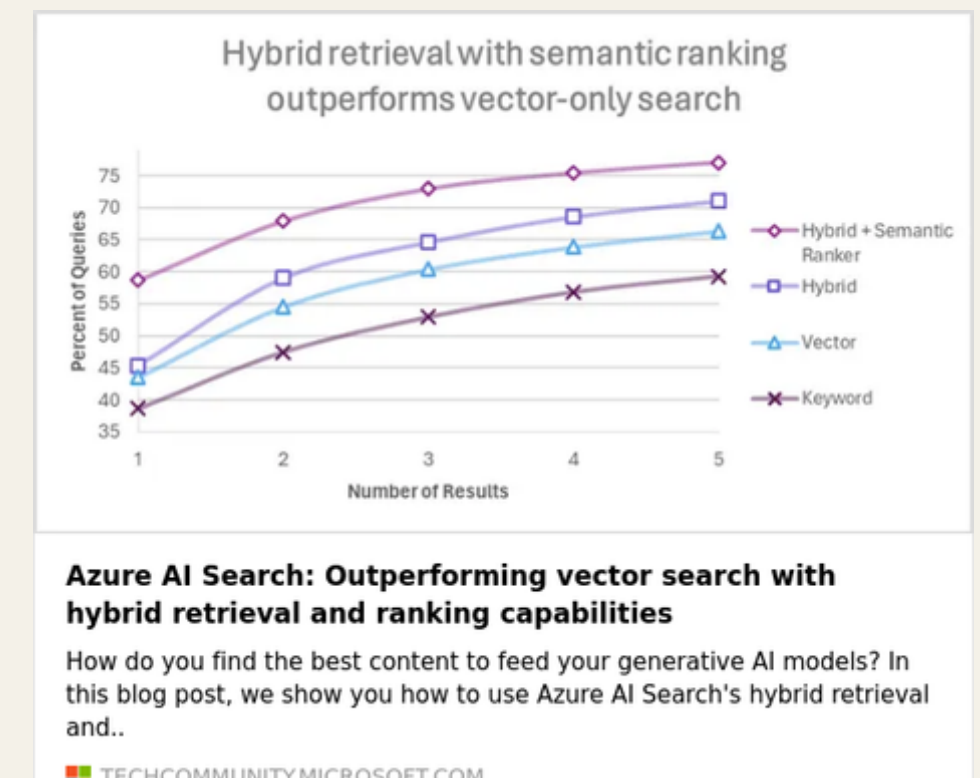
- 官網：不同領域會有不同最佳 Alpha (經驗談)
- 目標：教育領域 --> 政大選課相關 Dataset 做探究

參考文獻

IBM :
Blended RAG: Improving RAG
(Retriever-Augmented Generation)
Accuracy with Semantic Search and
Hybrid Query-Based Retrievers



Microsoft :
Azure AI Search: Outperforming vector
search with hybrid retrieval and ranking
capabilities

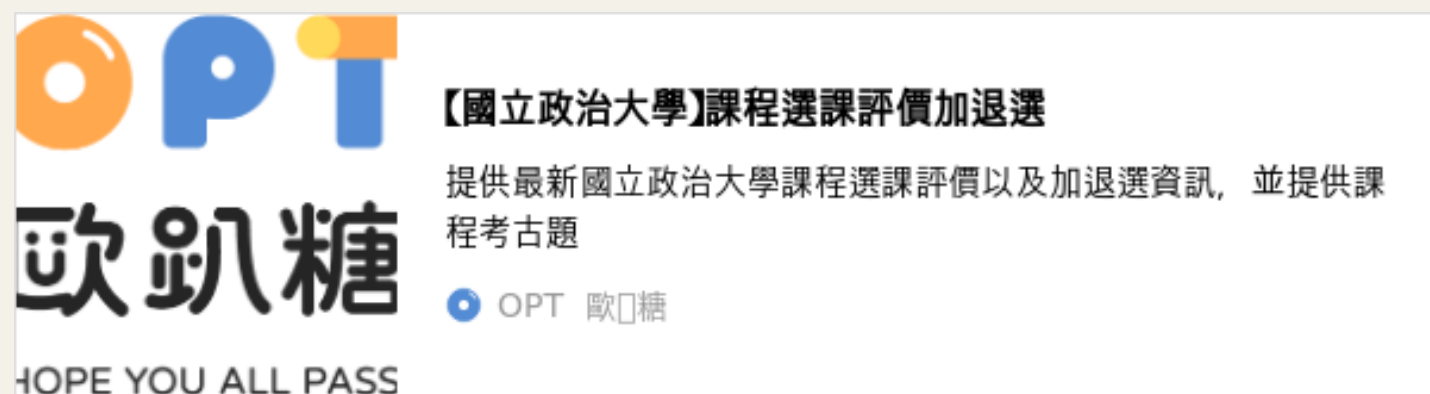


研究方法

1. 結合關鍵字匹配和向量語意搜尋
2. 30 題關鍵字形式 QA 測資 & 30 題語句形式 QA 測資
3. 利用 LLM 對測試資料進行自動測試和評分
4. 匯總不同權重配置（從 10:0 ~ 1:9）的測試結果
5. 將結果進行統計和視覺化，展示各種權重配置下的檢索準確性
6. 分析不同權重配置對檢索準確性的影響及造成此分布的原因

資料說明

- LLM 不會預先有的知識
- 教育領域：政大學生的選課需求
- 資料來源：OPT 歐趴糖

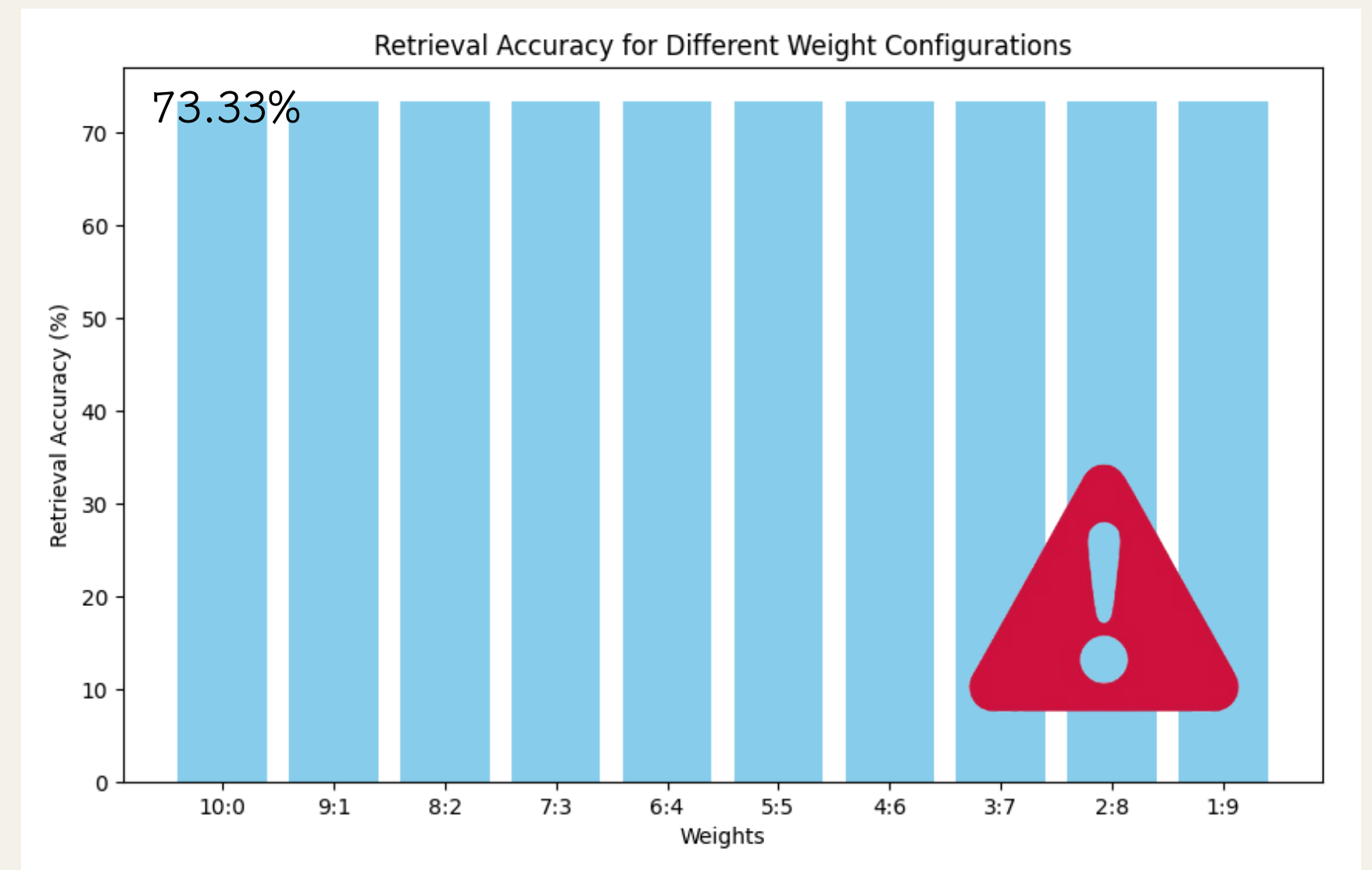
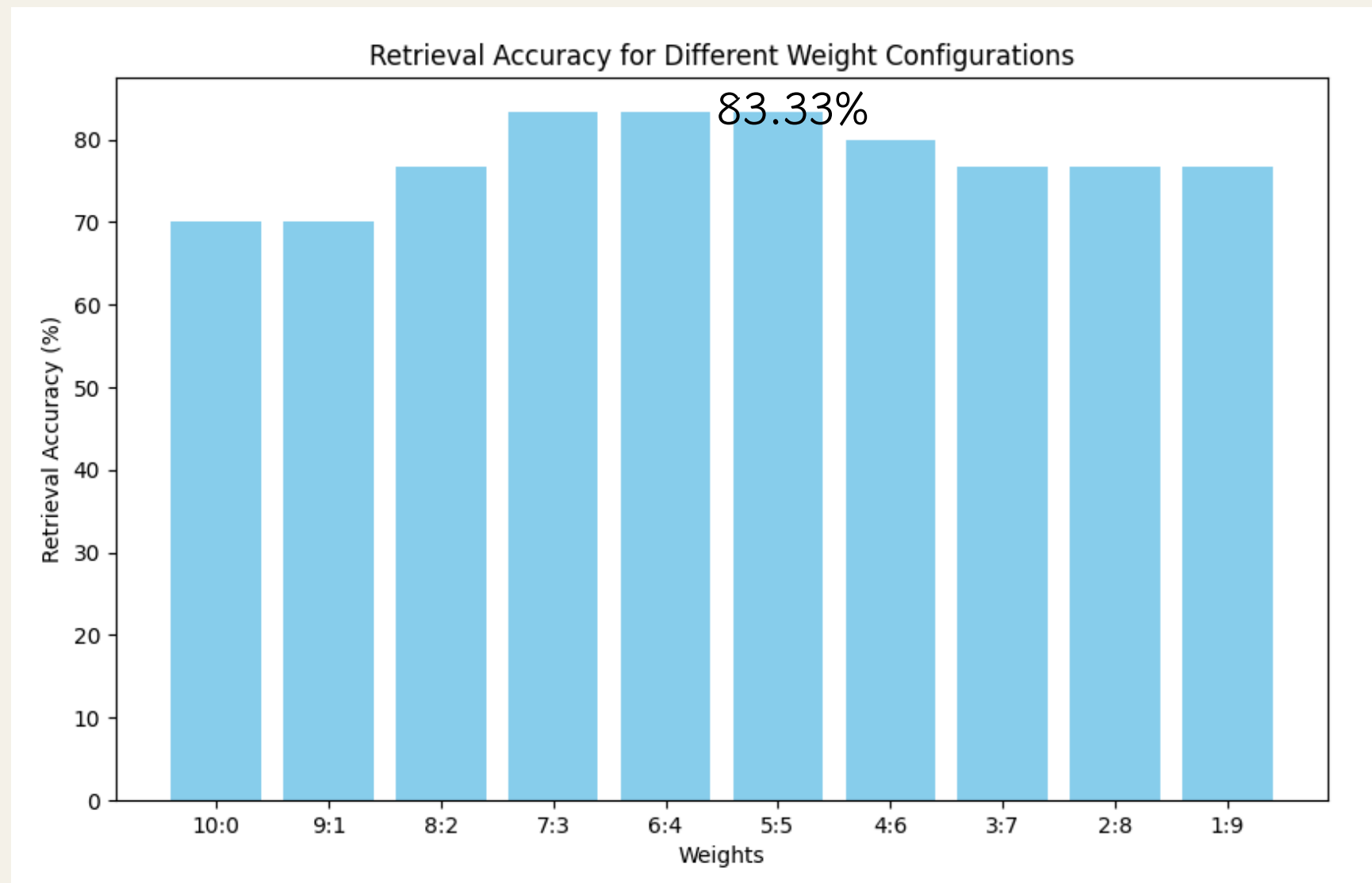


```
"title": "111-1 國立政治大學 林怡伶 程式設計一",
"description": "每堂課都有一個lab作業，老師會從頭開始教，沒基礎的不用擔心，會考小考和期中考，期末會分組做一個專題",
"evaluation": "輕鬆，分數高，老師人好"
JustinHsu
"title": "112-1 國立政治大學 陳志銘 數位圖書館與數位閱讀",
"description": "瞭解數位圖書館與數位典藏建置的相關議題，包括：數位圖書館的基本概念、發展趨勢、資源徵集與數位保存",
"evaluation": "老師會遲到或超時，作業難，討論很多"
JustinHsu
"title": "112-1 國立政治大學 邱炯友 數位學術出版與應用",
"description": "學術出版、學術期刊同儕評閱與創新、Doi及Ocid、掠夺性期刊、OpenAccess、政府資訊、檔案出版與開放",
"evaluation": "上課會想睡覺，作業蠻多，助教很熱情"
JustinHsu
"title": "112-1 國立政治大學 林巧敏 圖書館管理",
"description": "建立學生對於圖書館管理內涵的瞭解，並藉由主題報告與討論，深化學生對於各項議題的反思與解決問題能力",
"evaluation": "老師超好，作業壓力小，舒服自在"
JustinHsu
"title": "112-1 國立政治大學 羅崇銘 資料探勘",
"description": "PythonAI-機器學習、PythonAI-Azure視覺/語音/文字，簡單的程式設計。 其中有七次的非同步課程是要看",
"evaluation": "老師很愛聊天，同步課比較輕鬆，不用實際程式設計"
JustinHsu
"title": "112-1 國立政治大學 董祥開 基礎統計學",
"description": "大致在教社會的統計還有行政相關的統計資訊以便我們來管理及連接社會上許多相關的議題，並且增加我們對",
"evaluation": "我超餓我要吃雞排我要錢，我很餓，我超級餓"
JustinHsu
"title": "111-2 國立政治大學 陳陸輝 民意與調查",
"description": "本課程目的在使修習者對當前廣為運用的民意調查的研究方法有初步的了解，課程的內容涵蓋民意的理論與調",
"evaluation": "課程無聊，上課重點不明，學的東西沒用"
JustinHsu
"title": "111-2 國立政治大學 蔡惠名 實用台灣閩南語",
"description": "原則上真的是在教大家「學台語」，從台羅拼音到台語漢字。 老師會融入各樣教材給我們看，也會將語言的傳",
"evaluation": "臺灣閩南語，檢定考過有獎金，老師用心課程有趣"
```

程式碼 + 研究結果展示 (第一次試驗)

關鍵字形式 30 題 測試結果

語句形式 30 題 測試結果

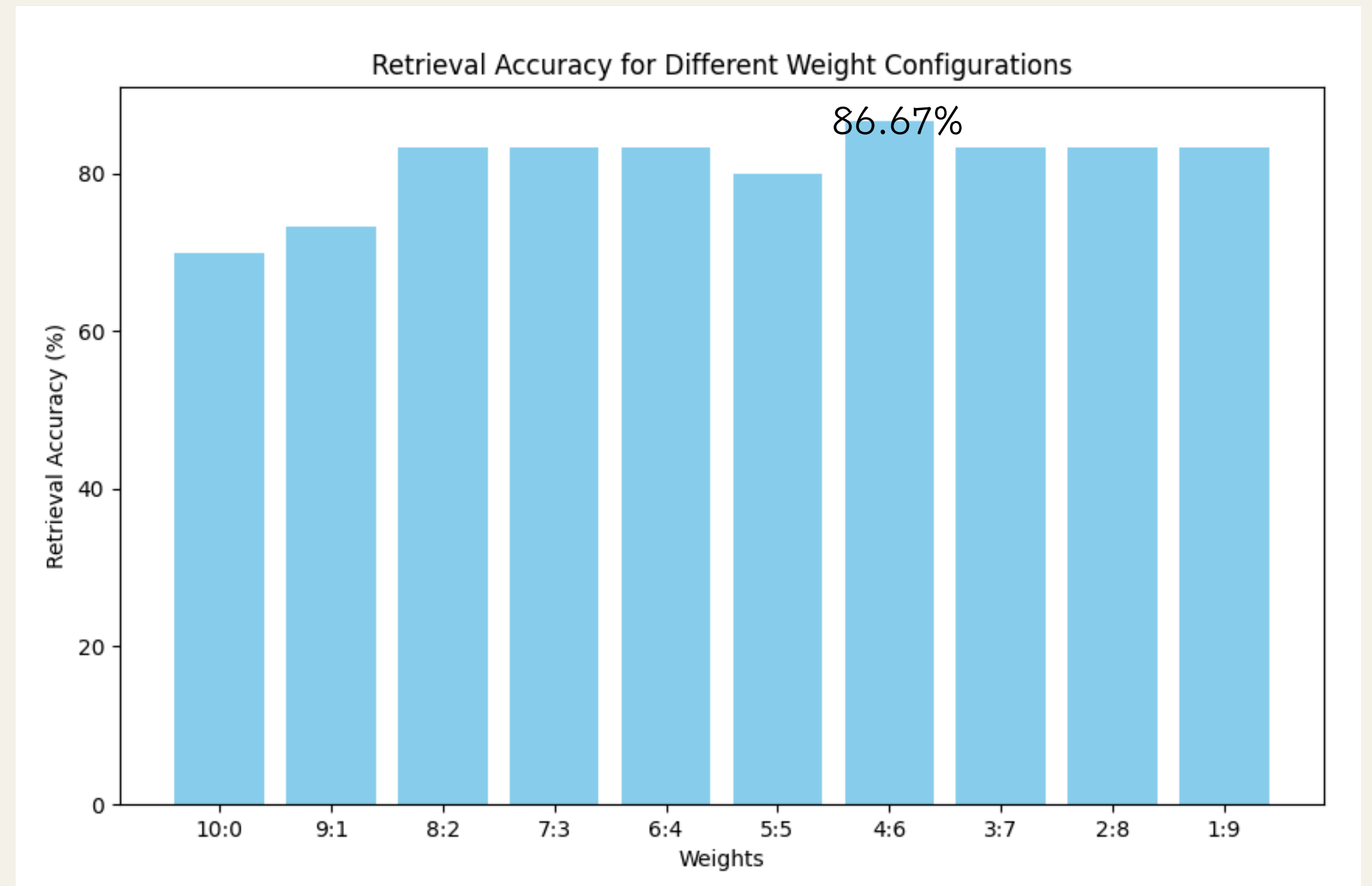


程式碼 + 研究結果展示 (第二次試驗)

改進 語句形式 30 題 測試方法

將此測試資料改為
結合 關鍵字 及 完整語句

- 使用 LLM 分詞



結論

第一次試驗

關鍵字 30 題 準確率分布

- 關鍵字組合起來也可做語意檢索，相當於整合語意與關鍵字的優勢

語句 30 題 準確率分布

- Weaviate 關鍵字搜尋分詞器僅支援英文，關鍵字評分過低，對結果無影響

結論

第二次試驗

語句 + 關鍵字 30 題 準確率分布

- 樣本太小 (僅 30 題有些特殊狀況會被放大)
- 向量搜尋可能會錯誤地優先考慮不相關的語意關聯，從而降低準確度
- 關鍵字經由 LLM 漂亮分好，可準確檢索

未來展望

1. 優化測試資料製作嚴謹性：

- 參考 Alpaca 等大型研究 QA 測試資料

<https://crfm.stanford.edu/2023/03/13/alpaca.html>

2. 加入新的變因：HyDE Search

<https://arxiv.org/html/2404.01037v1>

3. Finetune 提升準確率 (RAG + Finetune)

4. 與兩階段檢索方法的準確率比較

- 先做關鍵字搜索取 TopK，再做語意搜索取 TopN