

基於LLM的新聞不實性分析與段落查核系統

資訊系116級 陳青渝

指導教授：李政德

目的與動機

現今使用者多透過入口網站、社群平台或新聞首頁快速接收大量資訊，但新聞標題常可能包含誇張語氣、強烈因果、驚人數字或未充分說明的背景脈絡，使使用者在短時間內難以判斷內容是否需要進一步查證。

傳統事實查核通常需要人工搜尋來源、比對資料並判斷證據立場，流程較耗時，也不容易直接融入一般使用者的新聞閱讀過程。

特色

與傳統需要人工逐一搜尋與比對資料的查核方式相比，本系統將查核流程整合進使用者的新聞閱讀情境中。

系統不是單純讓 LLM 直接判斷真假，而是結合外部 evidence search、連結篩選與 stance detection，使查核結果更可信。

系統提供快速、智慧與精準三種模式，讓使用者可以依照速度與查核完整性的需求選擇不同流程。

系統功能

- 分為三個功能階段：
 1. 標題風險初篩
 2. 段落風險分析
 3. 段落連結查核

功能一：標題風險初篩

用於新聞首頁，只根據標題文字判斷該標題是否可能包含誇張語氣、驚人數字、強烈因果、重大健康風險或公共議題。

Step 1：前端擷取目前頁面可見標題

Step 2：後端過濾空標題，並限制最多 25 個標題

Step 3：將標題列表送入 LLM 進行風險初篩

Step 4：模型根據標題語氣、數字、因果、指控與高影響議題給出風險分數

Step 5：後端將風險分數轉換成低、中、高風險等級

Step 6：前端將風險分數顯示在新聞標題旁

功能一：標題風險初篩

首頁 新聞 股市 運動 TV 汽機車 購物中心 拍賣 氣象 App 下載

yahoo!新聞 新聞搜尋 登入

焦點 即時 娛樂影劇 國際 政治 社會地方 財經 運動 玩樂 品味 健康 遊

總代理  **和泰汽車**

13,900元 



- 豪雨不斷 高雄5區周三停... ● 0
- 熱顯像儀偵出人形！男奇... ● 30
- 台幣走弱 匯市爆逾1488... ● 30
- 梅雨狂下 脫離旱象「還... ● 30
- 149元「兩天救星」不怕... ● 0
- 賣光台積電 黃鐙輝換股... ● 30
- 鄰居氣爆致夫妻雙亡 初... ● 60
- 前港姐死於乳癌 醫示警... ● 30
- 預防雨季大樓供電中斷 ... ● 0
- 長時間久坐，美業人員面臨

(Dasuit 大適@坐墊) | Ad

yahoo!購物中心

\$1,540

NIKE 洞洞鞋 全面84折

Yahoo TV 新聞台

分析中

FakeLens

FakeLens

先看標題風險，再看段落理由，最後才做連結查核。

#1 [鄰居氣爆致夫妻雙亡 初判原因出爐](#) 60

涉及夫妻雙亡事件，且提到初判原因，需查證。

#2 [高金涉貪案核心 助理一千萬交保](#) 60

涉及貪污案，需查證的重大事實主張。

#3 [熱顯像儀偵出人形！男奇蹟獲救](#) 30

提到「奇蹟獲救」，可能帶有誇張情緒。

#4 [台幣走弱 匯市爆逾1488億元](#) 30

功能二：段落風險分析

從整篇新聞中找出值得進一步查證的段落，並標示其可疑原因與後續可能使用的查核 query。

Step 1：前端擷取新聞標題、網址、全文與每一段原文

Step 2：後端檢查文章長度是否足夠

Step 3：將整篇文章送入 LLM，抽取可疑段落

Step 4：模型輸出 text、score、reason、queries

Step 5：系統清理 query，避免模型補出錯誤年份或數字

Step 6：根據 threshold 過濾低分段落

Step 7：依照 score 排序，最多保留 max_chunks 個段落

Step 8：將抽出的 claim 對應回原始 paragraph_id

Step 9：前端在新聞頁面插入段落風險卡片

功能二：段落風險分析

< yahoo!新聞



打開App

焦點

即時

娛樂影劇

國際

政治

社會地方

財經

運動

玩樂

品味

健康

遊戲3C

AI 事件整理

台電台中區處長許墩貴表示，汛期來臨前台電均會進行相關巡檢作業，包括：檢視電線桿強度，如桿身有無龜裂或腐化、電桿基座的土基有無流失、電線有無鬆弛或脫落情形，以及支撐電線桿的支線部分是否穩固等，並同時進行配電線路及設備的全面檢視，防止颱風侵襲時造成傾斜或倒塌。此外，配電線路旁的樹（竹）木經常是造成線路故障的主因之一，都得事先編定計畫予以剪修整理。至於台中和平山區，台電台中區處在汛期來臨前，也已將相關器材設備，提前打包好運送就定位，避免屆時因道路中斷而發生的搶修器材匱乏狀況！

⚠ FakeLens 段落風險提示

0.60

可疑原因：需要確認台電是否在汛期前進行這些具體的巡檢作業。

抽出的主張：台電台中區處長許墩貴表示，汛期來臨前台電均會進行相關巡檢作業，包括：檢視電線桿強度，如桿身有無龜裂或腐化、電桿基座的土基有無流失、電線有無鬆弛或脫落情形，以及支撐電線桿的支線部分是否穩固等，並同時進行配電線路及設備的全面檢視，防止颱風侵襲時造成傾斜或倒塌。

查核這段

忽略

尚未進行連結查核

廣告

 Yahoo股市個股反查ETF
秒搜高含積量標的！
便宜上車去 

FakeLens



FakeLens

先看標題風險，再看段落理由，最後才做連結查核。

目前頁面

查核結果

說明

本文初步可疑度

0.70

高風險

本文初步可疑度為中；已標出 2 個高風險段落。此階段尚未搜尋佐證連結。

已標出 2 個段落；略過 1 個低於門檻的可疑段落。

高風險段落

#1 高風險段落提示

0.70

因滯留鋒面及西南氣流影響，中央氣象署今(9)日持續發布豪雨特報，昨日傍晚5點，也啟動了大規模劇烈豪雨事件加強作業；

可疑原因：需要確認中央氣象署是否發布豪雨特報及啟動大規模劇烈豪雨事件加強作業。

功能三：段落連結查核

當使用者在文章頁面看到段落風險卡片後，可以選擇按下「查核這段」，系統才會針對該段主張進行外部 evidence 搜尋與判定。

Step 1：使用者按下查核，前端送出 claim、internal_score、reason、queries、paragraph_id 與查核模式。

Step 2：後端建立 job_id，將查核工作放到背景執行。

Step 3：前端根據 job_id 持續查詢目前查核進度。

Step 4：後端根據 preview_mode 決定使用 none、smart 或 full 模式。

Step 5：系統根據 claim 與 queries 建立搜尋 query。

Step 6：系統搜尋查核網站、官方來源與一般網頁，取得候選 evidence。

Step 7：系統依照不同模式篩選候選連結。

Step 8：系統根據保留下來的 evidence 判斷 stance。

Step 9：系統將 stance 轉成 risk_score、label、reason 與 evidence links。

Step 10：前端將查核結果顯示在原本的段落風險卡片下方。

功能三：段落連結查核

加就是家庭主婦也沒關係，家庭主婦同樣一個月至少5,000，5,000元給家庭主婦去領」。

⚠ FakeLens 段落風險提示

0.80

可疑原因：賴清德提到的家庭主婦每月領取5,000元的政策需要查證其真實性和具體實施計畫。

抽出的主張：總統賴清德在520拋出「0到18歲成長津貼」政策引起熱議。他今（24）日說，有些人說，自己不是公務人員，也不是勞工，就是參加國民年金，「那就是家庭主婦也沒關係，家庭主婦同樣一個月至少5,000，5,000元給家庭主婦去領」。

查核這段

忽略

連結查核完成

連結查核結果：已有支持證據

查證風險 0.13

立場：supported | 多個證據來源確認賴清德提到家庭主婦參加國民年金可領取每月5000元的津貼，支持了該主張。

1. president.gov.tw/News/40060 官方 | gov.tw
2. 新生兒國家最高給108萬！賴清德再喊：未來國民年金「含家庭主婦在內...」 其他 | web
3. 六大社福津貼規劃調升 適時反映物價、加強照顧弱勢 官方 | gov.tw
4. 家庭主婦5000何時領？「一條件」就能月月領！大禮包整理：成長津貼最... 其他 | web

FakeLens UX 改良版

FakeLens

先看標題風險，再看段落理由，最後才做連結查核。

#1 已有支持證據

0.13

總統賴清德在520拋出「0到18歲成長津貼」政策引起熱議。他今（24）日說，有些人說，自己不是公務人員，也不是勞工，就是參加國民年金，「那就是家庭主婦也沒關係，家庭主婦同樣一個月至少5,000，5,000元給家庭主婦去領」。

立場：supported | 多個證據來源確認賴清德提到家庭主婦參加國民年金可領取每月5000元的津貼，支持了該主張。

原可疑原因：賴清德提到的家庭主婦每月領取5,000元的政策需要查證其真實性和具體實施計畫。

1. president.gov.tw/News/40060 官方 | gov.tw
2. 新生兒國家最高給108萬！賴清德再喊：未來國民年金「含家庭主婦在內...」 其他 | web
3. 六大社福津貼規劃調升 適時反映物價、加強照顧弱勢 官方 | gov.tw
4. 家庭主婦5000何時領？「一條件」就能月月領！大禮包整理：成長津貼最... 其他 | web

本網站使用相關技術提供更好的閱讀體驗，同時尊重使用者隱私，若您
意我們使用第三方 Cookie。若欲了解更多相關資訊，點這裡瞭解 隱私權政策

特價 食之香-酥脆派5入選



有片／朝柯文哲丟雞蛋畫面曝！北檢諭知8萬交保 「吃屎哥」找不到人辦保...

AviviD

查核模式	快速模式	智慧模式	精準模式
設計目標	追求最快回應速度	平衡速度與查核品質	追求較完整的查核流程
Query 使用量	較少，每個 claim 最多 2 個 query	中等，每個 claim 最多 4 個 query	使用段落分析階段產生的原始 queries
是否抓 preview	不抓	視 evidence 強度決定是否抓少量 preview	會抓取候選連結 preview
是否使用 GPT 過濾連結	不使用	使用，但只看 title、snippet、url、source	使用，並可搭配 preview 內容判斷
連結篩選方式	本地規則篩選	本地規則 + GPT 摘要層級過濾 + 選擇性 preview	多層搜尋 + preview + GPT 批次過濾
優點	速度最快、等待時間最短	速度與品質較平衡，是預設模式	查核流程最完整，判斷資訊最多
缺點	只看搜尋摘要，結果較保守	仍受搜尋結果與 query 品質影響	速度最慢，preview 可能受網站限制
適合情境	快速展示或初步查核	一般使用情境	需要較完整查核時使用

系統限制

1. 搜尋結果仍會受到 query 品質影響，若 query 沒有命中正確關鍵詞，可能找不到合適 evidence。
2. 部分網站可能無法順利抓取 preview，導致系統只能根據 title、snippet 與 url 判斷連結相關性。
3. 背景資料不一定能直接證明新聞事件本身，因此系統需要區分「背景佐證」與「直接事件佐證」。
4. 外語與專業文獻搜尋仍受限，若核心概念未正確轉換為英文專業關鍵詞，可能影響 evidence 的搜尋命中率。

未來展望

1. 未來可加入更多可信來源，例如官方資料庫、醫學文獻資料庫與國際查核網站，提高 evidence 覆蓋率。
2. 未來可針對不同新聞類型設計更精準的 query 生成規則，例如醫療、政策、國際事件與財經新聞分開處理。
3. 未來可建立資料集，用來評估段落風險分數、連結篩選結果與 stance 判定準確率。