

# << 社群媒體分析 >>

指導教授:黃三益 教授

第三次讀書會報告\_第十四組

一、分析主題：川普 2025 年關稅政策對台灣產業、股市與就業輿論的影響

二、組員名單：

馮蕙芳\_N124080003

高健芝\_N124080006

廖胤翔\_M134610022

Deepan\_M134610032

孫郁琪\_M136020020

胡賀\_J136020007

三、分析工具：• 中山大學工作流程平台 Tarflow • 工作流程名稱：讀書會(三)

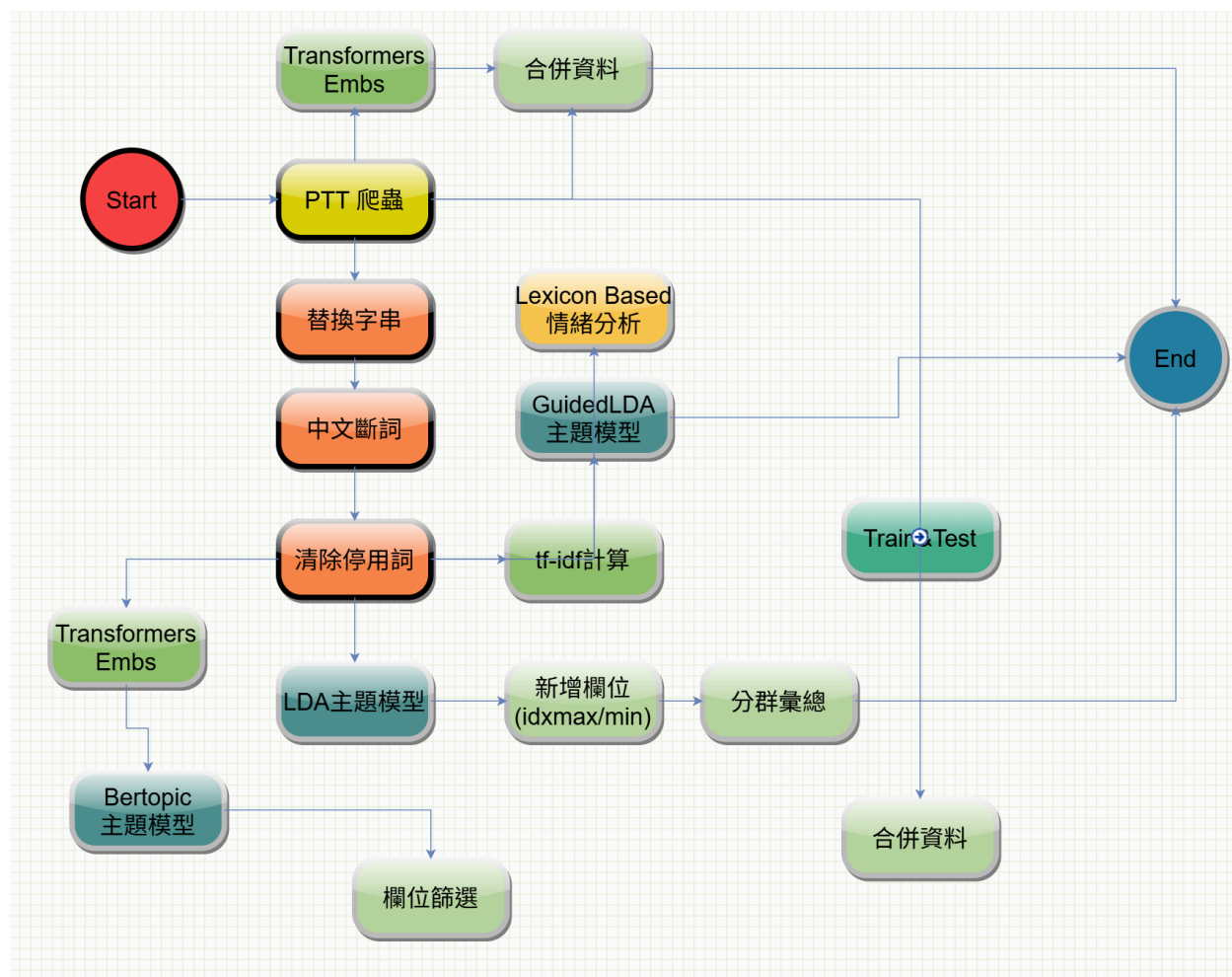
四、動機：2025 年 4 月，美國總統川普宣布對多國恢復高額關稅，主張製造業回流美國。這項貿易政策震撼全球市場，不僅引發美股波動，也可能深刻影響仰賴出口與科技製造的台灣。而我們這組關注的是：台灣民眾與市場參與者如何在社群平台上回應這波政策變動？是否反映出對投資信心、就業穩定、產業未來的集體情緒與判斷？

因此，我們決定以 PTT 為資料來源，針對與金融、產業、政治討論相關的六個看板進行資料蒐集與文字分析，觀察政策從發酵到實施期間，台灣社群的輿論焦點、認知變化與集體情緒。

透過這次研究，我們希望能理解：一項國際經濟政策，如何在台灣引發跨產業與跨領域的社群討論與反應，並從中看見公共政策與民間輿論間的動態關係。

五、資料來源：PTT 上 Stock(股票)、Finance(財經)、ForeignEX(外匯)、Tech\_Job(科技業)、soft\_job(軟體工程)與 HatePolitics(政黑)共6看板。

工作流程:



# 一、資料預處理

## PTT 爬蟲 (4)

參數設定

任務結果

選擇看板 \*

Finance(金融業)

Food(美食)

ForeignEX(外匯)

forsale(二手買賣)

gay(甲)

GetMarry(結婚)

Gossiping(八卦)

搜尋關鍵字 ⓘ

川普  
關稅  
貿易戰  
製造業  
產業外移

排除關鍵字 ⓘ

以換行區隔，e.g.  
壽山動物園  
猴子  
...

搜尋起始日期

2023/10/15

搜尋結束日期

2025/05/10

儲存更改

為了準確捕捉台灣社群對此政策的反應，我們以 2023 年 10 月 15 日至 2025 年 5 月 10 日為資料蒐集期間。選擇此段時間的原因，是因為川普在 2024 年總統初選與競選活動期間，已開始強烈主張「製造業回流」與「對多國徵收高額關稅」，而在 2025 年 4 月正式提出具體貿易政策後，台灣媒體與論壇上出現大量討論。我們希望涵蓋政策初次提出前的預期反應與政策發佈後的實際回應，因此延長觀察期間以完整掌握討論演變。

本研究使用 TarFlow 爬蟲系統針對 PTT 六個看板進行資料蒐集，包括 Stock(股票)、Finance(財經)、ForeignEX(外匯)、Tech\_Job(科技業)、soft\_job(軟體工程)與 HatePolitics(政黑)。這些看板反映不同使用者社群的視角，從資本市場、就業結構、產業調整到政治立場與國際關係，皆有代表性。

在搜尋設定方面，我們設定了與主題高度相關的關鍵詞作為資料過濾依據，包括：

川普 關稅 貿易戰 製造業 產業外移 製造回流 台灣製造 出口 投資 股市 科技股 AI股 半導體 供應鏈 產業鏈 外資 資金 裁員 就業 失業 房價 貸款 升息 通膨 經濟 經濟風險 消費信心 工程師 簽證 矽谷。

## ≡ 替換字串 (28)



參數設定

Input - 4

任務結果

選擇處理欄位 \*

artContent

替換字串設定 ⓘ

\n\n>> °  
\n>> ,  
Sent from w\|+>>  
https>>  
com>>

選擇替換規則檔案 ⓘ

-----請選擇-----

儲存更改

接著，我們簡單進行了替換字串的設定並進行中文斷詞。為了強化模型辨識與本研究主題相關的重要語意，我們根據研究問題手動設定了一組自定義詞彙，並分配相對應權重，以提升模型在主題聚類與嵌入表現時的指向性與辨識精度。

這些詞彙如「美國」「關稅」「貿易戰」「中國」「台灣」等，皆為川普2025年關稅政策中，台灣社群可能特別關注的核心焦點。

## ≡ 中文斷詞 (7)

參數有做更動，建議重新執行



參數設定

Input - 28

任務結果

選擇處理欄位 \*

result

定義詞彙 ⓘ

美國 1000  
關稅 1000  
貿易戰 1000  
中國 1000  
台灣 1000

選取字典 ⓘ

-----請選擇-----

儲存更改

接著，我們自定義了一些停止詞。

## 清除停用詞 (8) 參數有做更動，建議重新執行

參數設定

Input - 7

任務結果

是否清除單字元

是

清除英文字母

否

清除換行符號

是

清除html tag

是

是否轉為小寫英文

是

清除數字

是

清除特殊標點符號

是

自定義停止詞

可以  
就是  
仍然  
沒有  
因此  
以及

儲存更改

## Transformers Embs (9) 參數有做更動，建議重新執行

參數設定

Input - 8

任務結果

選擇處理欄位

result

模型名稱

distiluse-base-multilingual-cased

儲存更改

## Transformers Embs (9) 參數有做更動，建議重新執行

參數設定

Input - 8

任務結果

### 任務結果

Show 10 entries Search:

system_id	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0.051125	0.036930	-0.008493	-0.032803	-0.000912	-0.076561	-0.012317	0.039398	0.034910	0.006757	-0.014134	0.014319	0.016481	-0.009463	-0.000912
2	-0.024193	0.049735	0.030413	-0.018597	-0.044104	-0.018961	-0.004623	0.037304	0.044952	0.021302	-0.021663	-0.004290	0.012307	0.030251	-0.000912
3	0.024756	0.034448	-0.020001	-0.009419	0.027766	-0.032691	-0.020133	0.053820	-0.034829	0.031723	-0.018398	-0.007840	-0.053120	-0.026397	-0.000912
4	0.008952	0.003989	0.002199	-0.006036	-0.024605	0.001566	0.002579	0.040912	0.004018	0.021872	-0.017281	-0.004552	0.009292	0.017833	-0.000912
5	0.004856	0.055898	0.003620	0.013779	0.046246	-0.068610	-0.037105	-0.015968	0.034428	0.004285	-0.024397	0.018627	-0.035369	-0.016002	-0.000912
6	0.002139	0.039825	0.000992	-0.019980	-0.058868	-0.016745	-0.006379	0.024895	0.012037	0.005980	-0.005464	-0.013244	-0.027719	0.035872	-0.000912
7	0.023806	-0.000432	0.025546	-0.007525	-0.007956	0.029331	0.013808	0.026080	0.057695	0.031799	-0.028517	-0.006392	-0.020982	0.025283	-0.000912

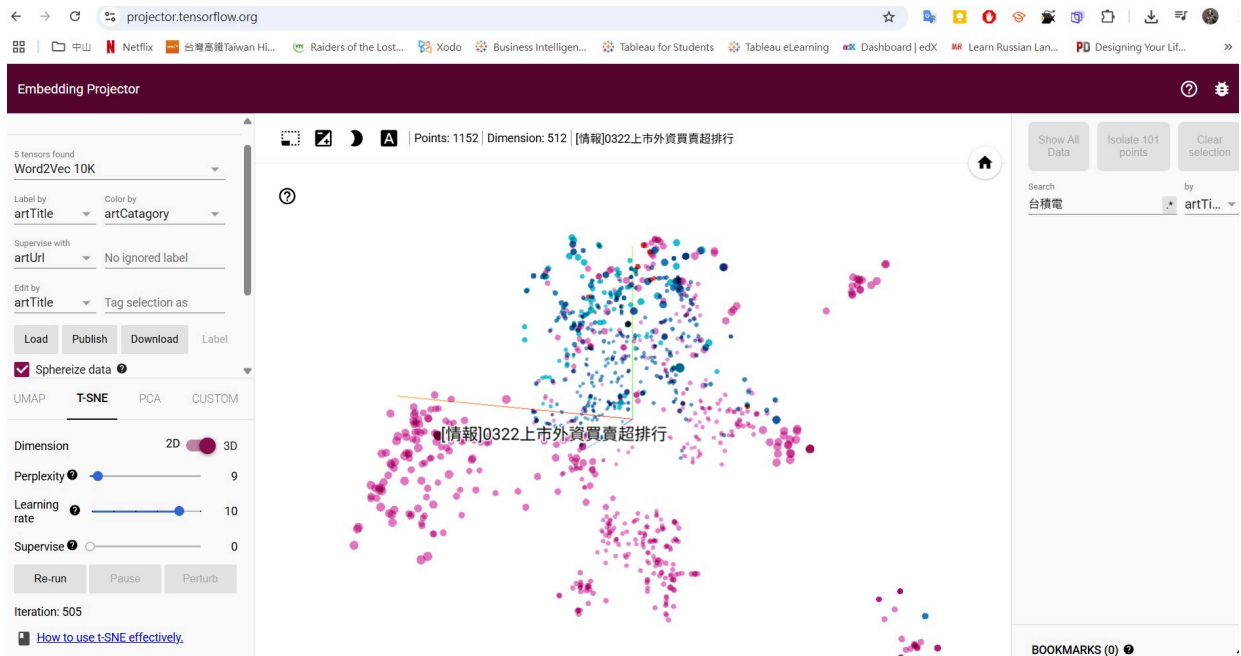
全螢幕瀏覽

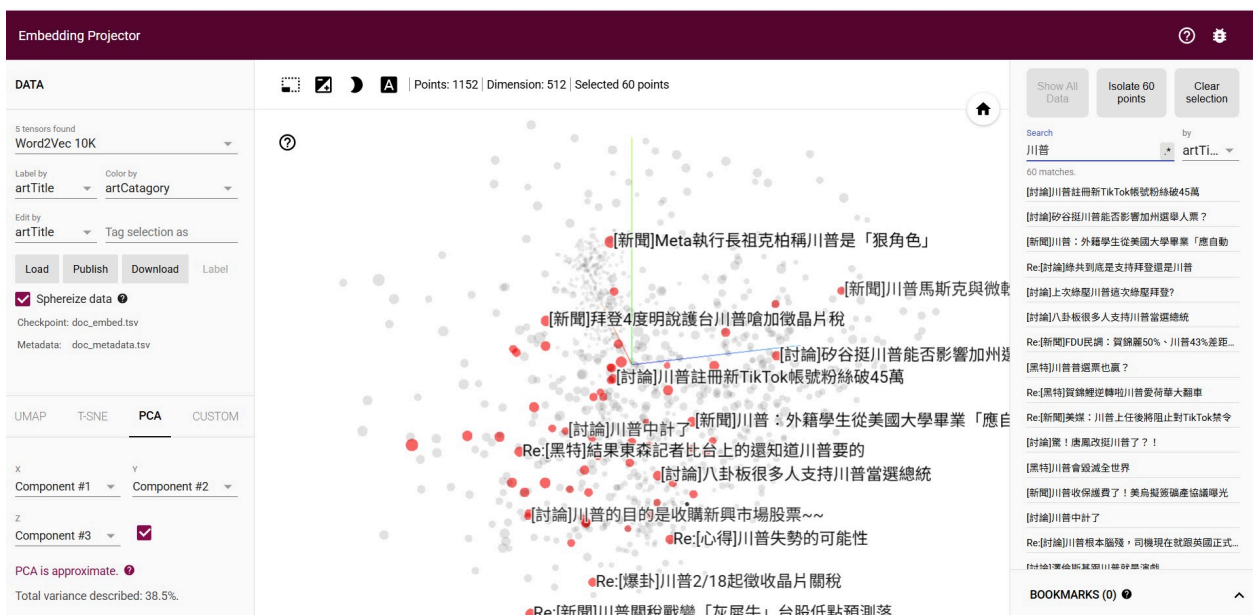
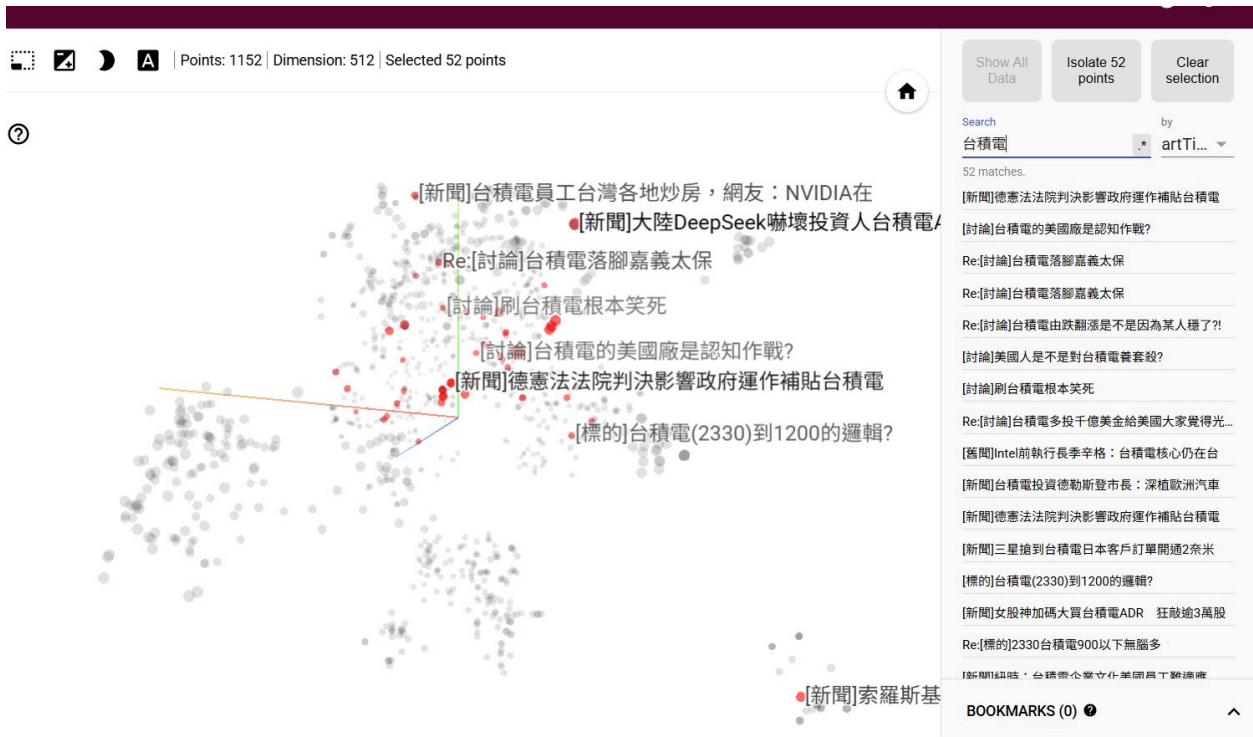
點我下載完整CSV資料

點我下載完整Rdata

點我下載完整json資料

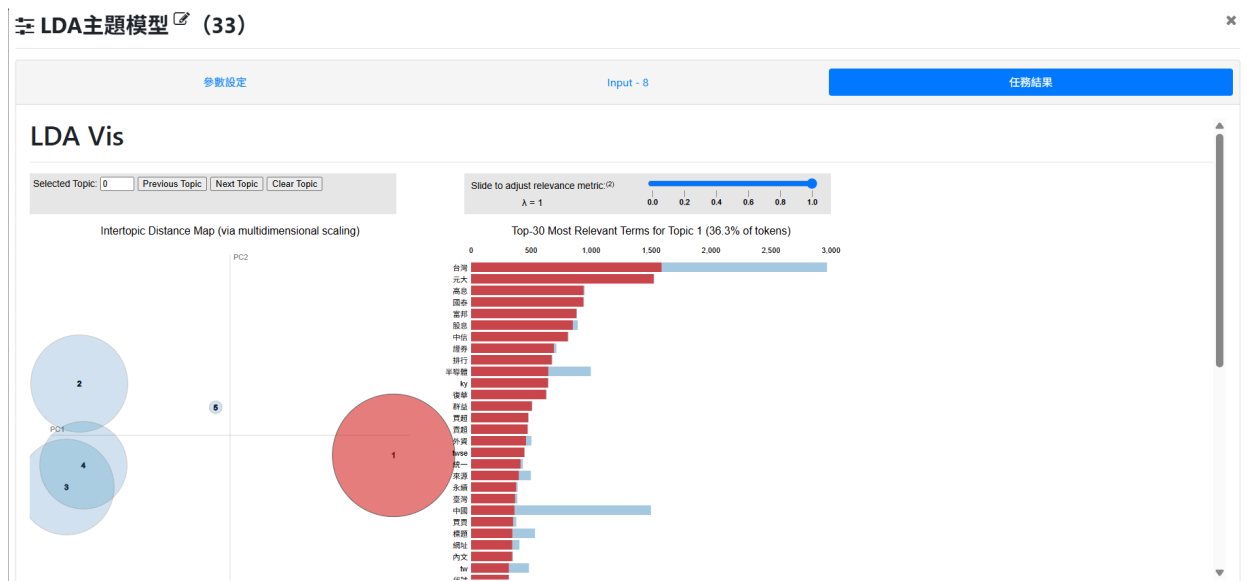
接著我們在設定以及跑完Transformers Embeddings後，將其結果下載成Csv檔案，並在整理欄位及元資料後，將其轉成Tsv檔案再使用Embedding Projector 進行視覺化呈現。





## 二、LDA模型

我們使用LDA 主題模型將文本分為五個主題，以下列出前三個主題的視覺化結果與內容解讀：



在主題一 (Topic 1) 中，模型擷取出最具代表性的關鍵詞包括「台灣」、「元大」、「國泰」、「聯邦」、「中信」、「高息」、「智能」、「半導體」、「排行」，最具代表性的關鍵詞明顯集中於「台灣金融投資產業」、「金融機構」、「半導體與高科技產業」、「財經數據排行與表現」等領域。這反映該主題在樣本文本中的代表性，即強調台灣地方與金融投資、半導體產業息息相關的資訊。



## ☰ LDA主題模型 (33)

在產業應用方面的關鍵詞：“公司”、“產品”、“invest(投資)”、“美國”、“中國”等詞出現，顯示AI產業的國際發展、企業參與與產品推出是討論重點。

在相關機構或企業的方面的關鍵詞：“openai”，這個字詞的出現顯示出該主題涵蓋了非營利組織OpenAI Inc 這間公司及其技術產品(如ChatGPT)。

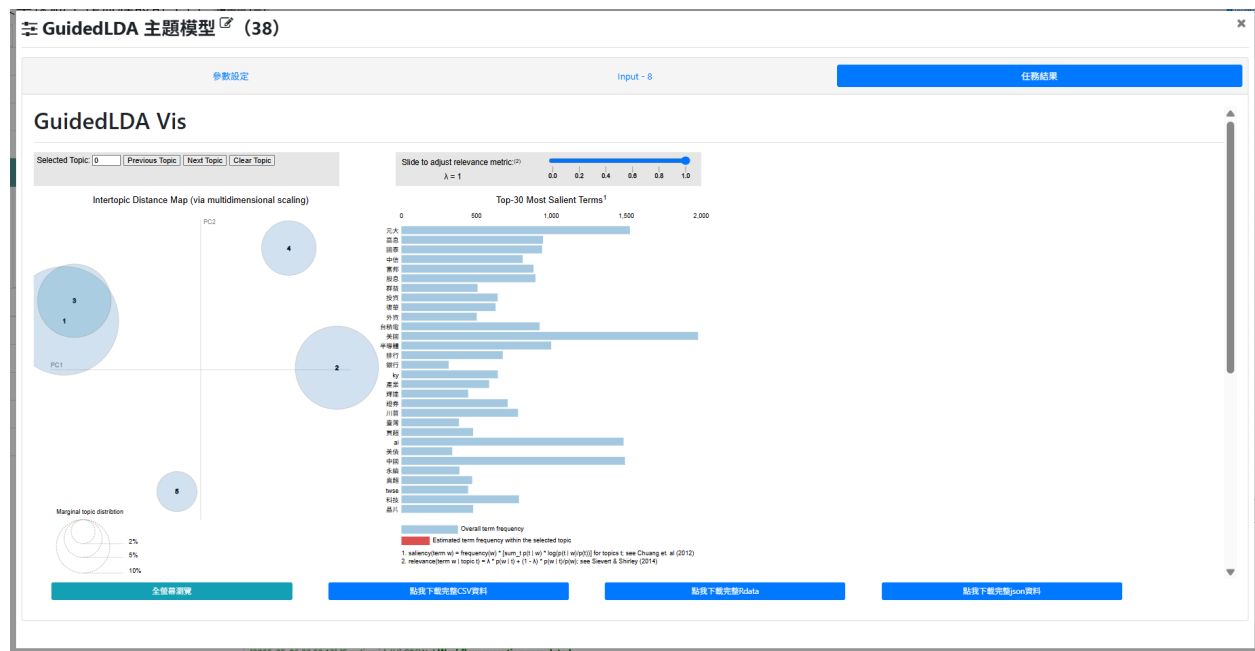
其次也包含像“未來”、“發展”、“meta”等字詞，這也暗示著這些文章的討論不僅僅侷限於現況，還有AI未來發展方向、趨勢及影響。

總合來看，我們設定了5個主題，從統計資料可看出文本內容可以分出多重關聯性話題，具備一定主題多樣性



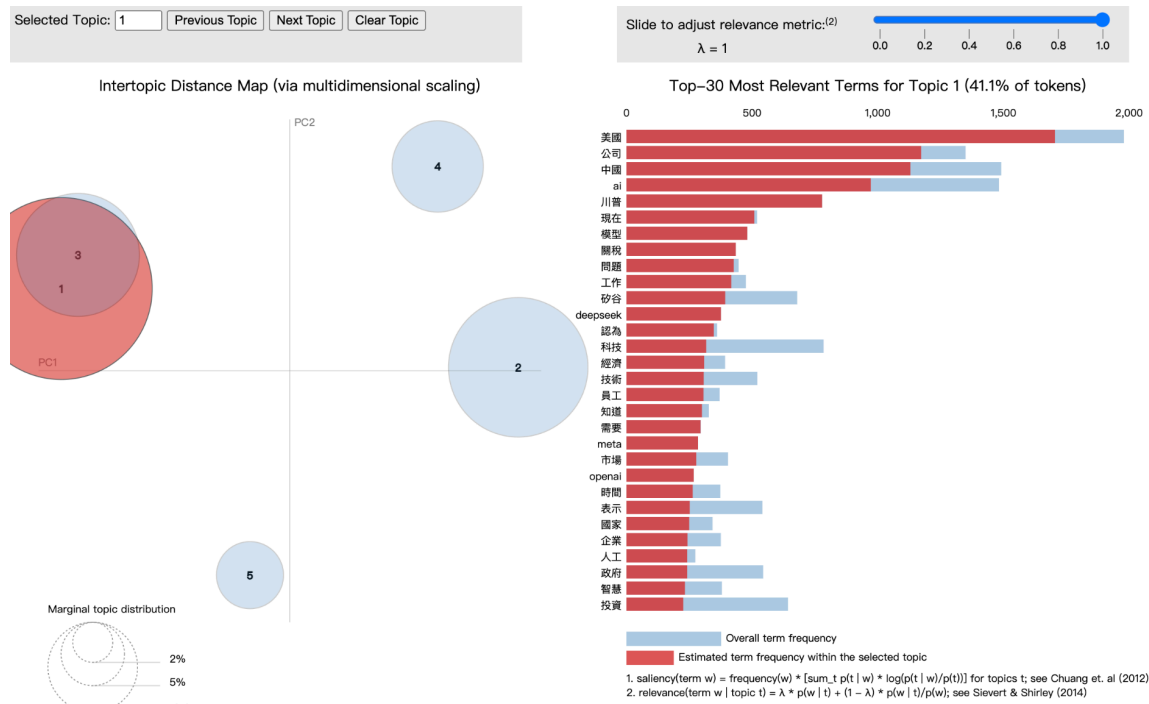
PMI: -1.482) 指數值偏低，表示主題一致性偏低，也暗示部分主題內的詞語關聯性不夠強，或者主題之間可能不夠清晰，另一方面也可能是因為文本字數偏少或主題間的關鍵詞重疊，導致模型無法清晰區分各主題特徵。在主題區別性 (Cv: 0.485) Cv指標接近0.5，反映5個主題之間在關鍵詞的區隔程度尚可，但仍有改進空間。在混淆度(Perplexity): (496.53) 高混淆度值代表模型對文本擬合程度尚可，但數值偏高可能暗示模型在主題推測上存在過擬合或文本資訊分布不均的問題。

因此我們進一步使用 GuidedLDA 模型



我們也有使用 GuidedLDA 模型，設定5個主題進行建模，並依據我們事前設計的種子字「美國, 關稅, 貿易戰, 中國, 台灣, 科技, 經濟, 矽谷, 產業, 股市, 金融, 外匯, 政治, 政策, 關稅政策, 進口稅, 生產成本, 供應鏈, 產地轉移, 科學園區, 產地設在美國, 順差, 逆差, 外銷訂單, 出口影響, 全球經濟不穩, 經濟風險, 經濟不穩, 通膨, 升息, 裁員, 就業不穩, 遠端工作, 海外就業, 外派, 美國缺工, 矽谷工程師, 人才流動, 房價會跌, 房貸, 買股, 股票保證金, 貸款壓力, 補繳, 風險, 保守, 投資, 保值, 防禦型配置, 投資型保單, 成本上升, AI股, NVIDIA, 晶片禁令, 晶片法案, 台積電, 台灣半導體」等等

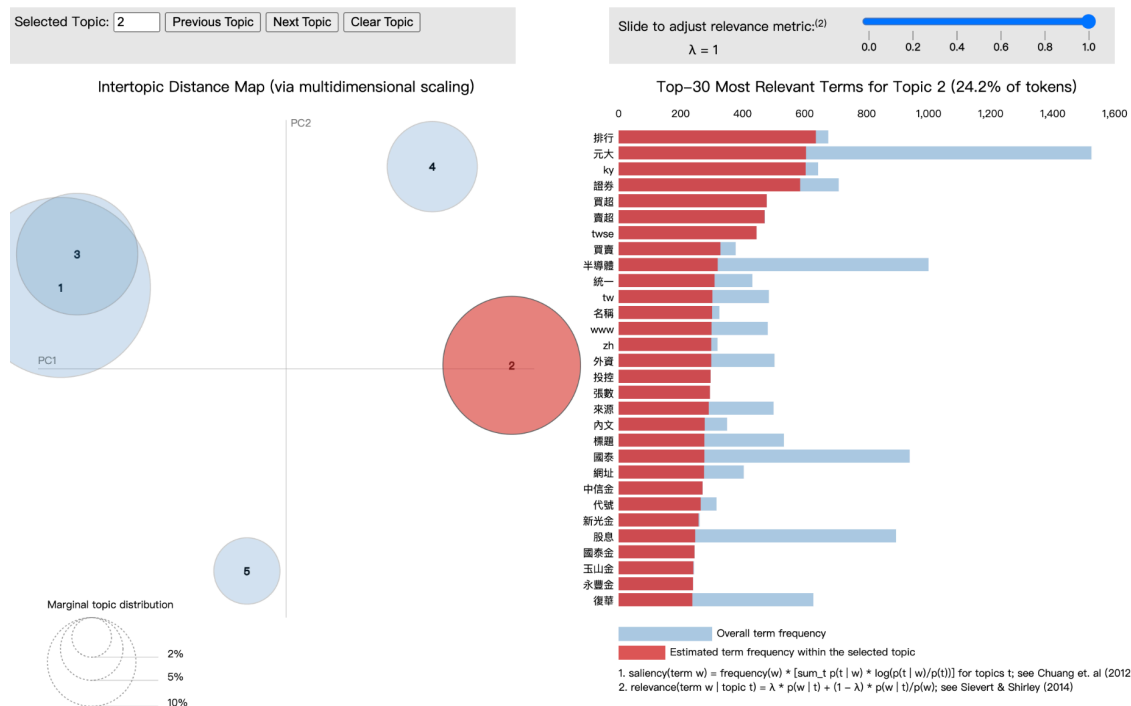
作為語意引導，確保模型能聚焦在事件相關的討論面向上。這樣的設計有助於提升主題的可解釋性，避免出現模糊或無意義的分類。



## 主題 1: 全球 AI 競逐與美中科技地緣政治

關鍵詞主要集中在「美國」「中國」「AI」「矽谷」「openai」「meta」「川普」「關稅」「政府」「市場」「投資」「人工智慧」「企業」……可看出在大家關注AI 全球競爭格局、科技巨頭與政治勢力，例如：「deepseek」與「矽谷」等新詞代表 AI 開源趨勢、「川普」「關稅」代表政策干預。

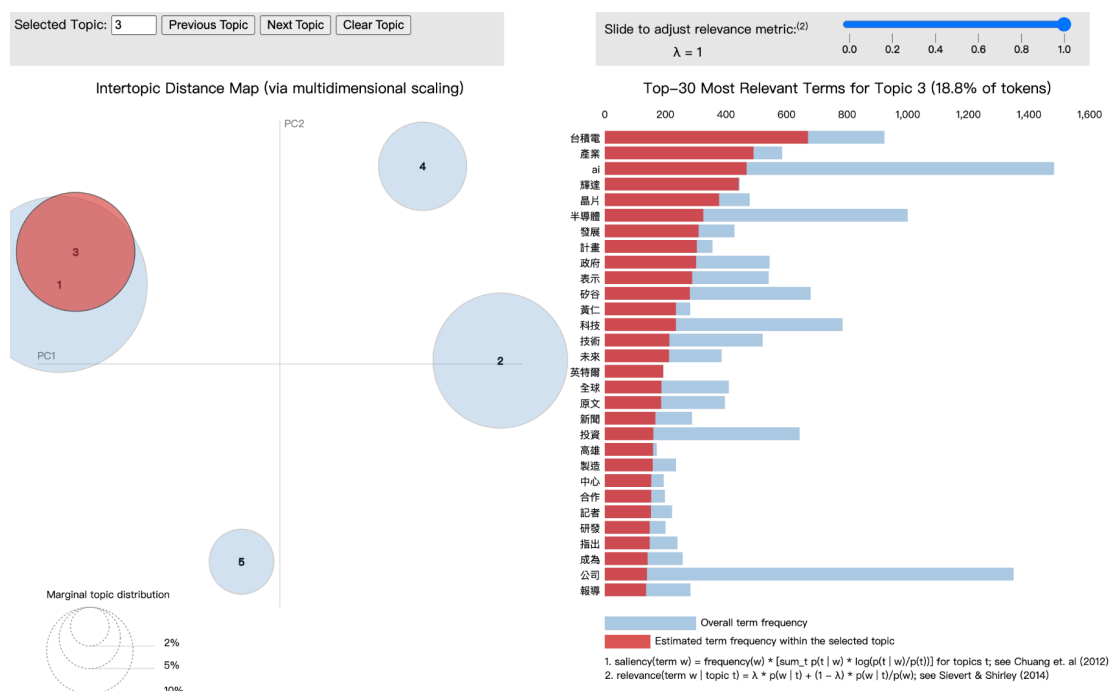
1. 企業與技術生態: OpenAI、Meta 及矽谷公司領軍的生成式 AI 浪潮。
2. 地緣政治與政策摩擦: 川普、關稅、政府、國家等詞揭示美中在關鍵技術與供應鏈管制上的緊張關係。
3. 技術擴散與管制: 一方面生成式 AI 進入產品化、商業化; 另一方面美中相互祭出晶片與雲端的出口管制。
4. 政策風險外部影響: 關稅及出口禁令不只影響半導體, 對雲端服務、資料跨境流動也形成合規壓力, 迫使企業重新評估供應鏈布局。



## 主題 2 | 台股資金流與金融控股動態

關鍵詞主要集中在「排行」「元大」「ky」「櫃買」「twse(證交所)」「買超／賣超」「外資」「半導體」「張數」「國泰金」「中信金」「新光金」「股息」「代號」，以此可以發現該主題聚焦在台灣股票市場的資金流動與法人操作，特別是外資與金控公司如何影響市場。

1. 市場交易指標:外資買賣超、張數、漲跌幅度等，高頻出現代表投資人情緒與市場流動性。
2. 金控／券商角色:元大、國泰、中信、新光、永豐、玉山等金融控股與券商頻繁被提及，反映其在研究報告、ETF 及資金調度上的影響力。
3. 半導體影響擴散:雖以金融詞彙為主，但「半導體」仍列入前十，呼應電子權值股對台股走勢的牽動。
4. 資金面主導台股情緒:外資動向與本土法人(投信／金控)互動，形成短線股價波動與 ETF 申購／贖回循環。
5. 金控業者積極推動 ETF、綠色金融商品並參與 AI、半導體等題材，形成資金與話題的導向中心。



### 主題 3 — 半導體產業擴張與 AI 供應鏈

「台積電」「AI」「晶片」「半導體」「投資」「高雄」「研發」「Jensen Huang (黃仁勳)」「英特爾」「合作」「製造」「政府」「中心」……可歸納為：

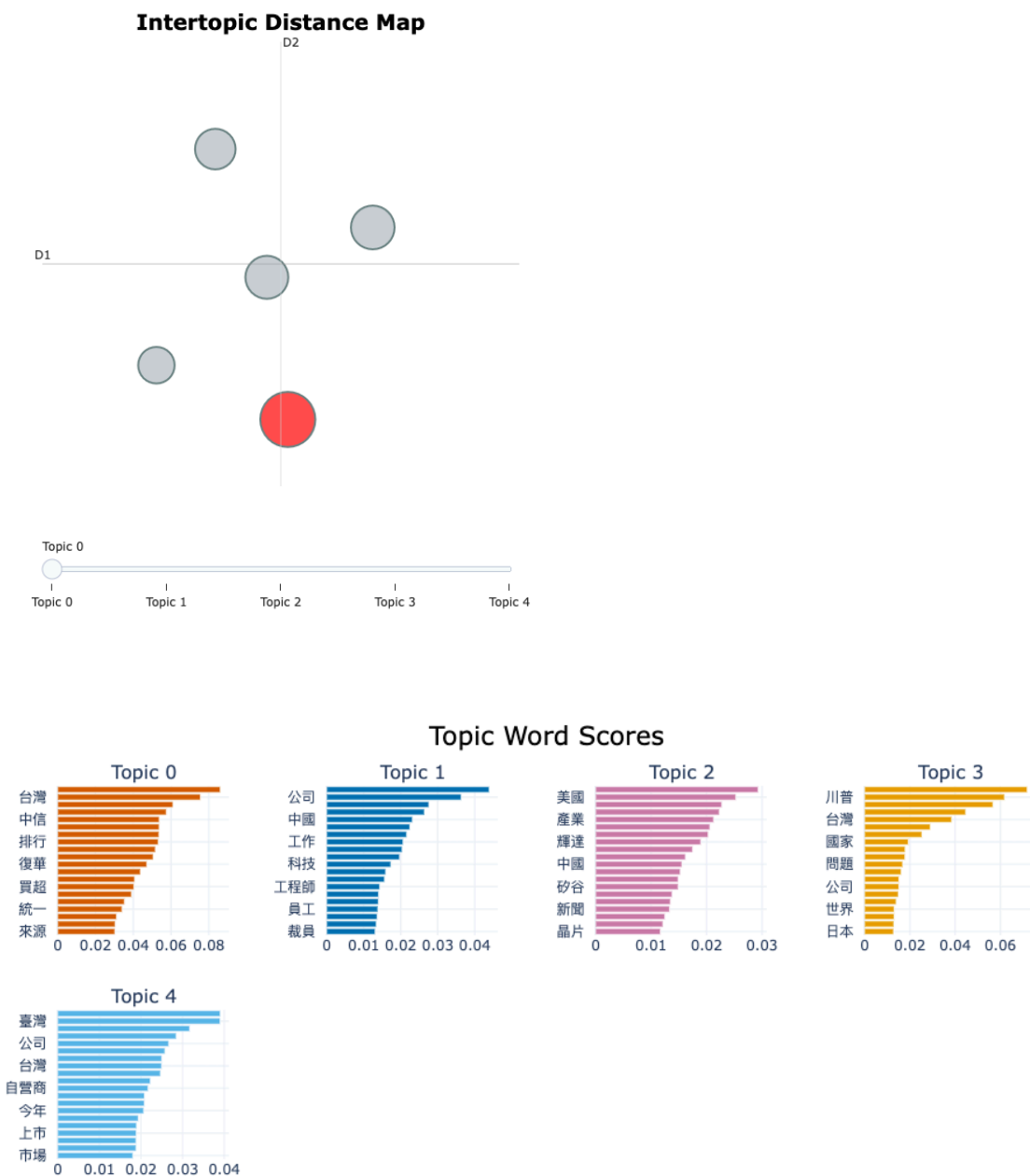
1. 先進製程競賽：台積電、英特爾、Nvidia (Jensen Huang) 等在 2 nm 以下製程、HPC／AI GPU 及 CoWoS 封裝技術上角力。
2. 在地化與擴廠：高雄地名隱含台灣南部科學園區新廠布局；「政府」「計畫」指出國家政策補貼與基礎設施配套。
3. 全球供應鏈協作：關鍵詞「合作」「研發」「中心」與美日歐在先進封裝、材料、IP 設計的結盟。

從主題分布比例來看，主題一(41.1%)佔比最高，顯示各個版留言中對AI技術發展與國際科技競爭的關注最為集中，尤其是美中關係、生成式AI、以及科技巨頭的政策與商業行動，成為推動輿論與新聞關注的主要來源。此一主題也反映出台灣作為供應鏈關鍵位置，在全球科技地緣政治中所處的微妙與高風險位置。主題三(18.8%)則與主題一語意距離最接近，兩者之間形成明顯的交集與重疊，顯示在談論AI技術競爭的同時，大眾也高度關注其對半導體製造與產業擴張的實質影響。例如台積電在南部擴廠、NVIDIA 與政府的政策合作、以及地方產業發展等面向，說明產業面的關注與科技政策間具有強烈聯動。

相較之下，主題二(24.2%)雖然在比例上次於主題一，但在視覺圖上與其他主題明顯區隔，說明此主題內容語意較為獨立，聚焦在台股資金流、法人買賣超與金控動態等投資面討論。這種資本市場與科技產業相互出現，凸顯出台灣媒體中，AI題材與半導體股已成為投資的焦點。

整體而言，前三大主題呈現出台灣社會當前對「AI 崛起下的國際風險」、「半導體產業鏈」以及「金融市場反應」等議題。從我們選中的六個PPT討論版，可挑出高頻出現的詞彙如「投資」「政府」「技術」「買超」「晶片」「台積電」「OpenAI」等，這些議題已廣泛滲透至政策、產業層面，且與國際政治經濟議題高度相關聯。

### 三、Bert Topic



總體來說, BERTopic分析的各個主題 (Topic 0–4) 分布於不同象限, 未大量重疊。代表主題之間具備一定的語意獨立性, 模型有良好區辨度。

BERTopic 的分析結果產生了五個主要主題:

Topic 0: 台灣市場與外資動向



- 代表詞:台灣、中信、排行、復甦、買超、統一、來源
- 討論台灣金融市場的投資情況, 特別是中信、統一等大型金融機構的表現。
- 關稅政策改變後, 外資對台灣市場的佈局調整, 可能影響市場的買超或賣超行為。

#### Topic 1: 美中貿易與科技產業影響

- 代表詞:美國、中國、新聞、晶片、表態
- 聚焦美中貿易戰背景下, 矽谷科技公司和晶片產業的調整, 台灣半導體產業可能面臨供應鏈挑戰。
- 討論矽谷、新聞報導和台灣出口產業的後續影響。

#### Topic 2: 科技人才市場與技術發展

- 代表詞:公司、deepseek、科技、工程師、人才、員工
- 關注美國政策變動後, 台灣是否能吸引更多國際科技人才, 或是面臨人才流失。
- 特別強調技術研發、工程師流動與公司技術創新。

#### Topic 3: 貿易政策與國際影響

- 代表詞:川普、台灣、國家、問題、貿易、通膨、日本
- 直接回應政策核心——川普復徵關稅, 社群討論其對全球與台灣的戰略經濟意涵。包括通膨、貨幣與進出口議題。
- PTT HatePolitics、Finance 是主要的討論平台。

#### Topic 4: 台灣股市與投資信心

- 代表詞:投資、公司、股價、台灣、今年、股息
- 關注台灣股市在政策改變後的市場波動, 投資人對股價走勢、股息政策的信心變化。
- 主要出現在 PTT Stock、Finance, 特別是市場波動時的集體情緒反應。

### 政策影響的初步觀察

根據 BERTopic 的結果, 可以觀察到以下幾點:

1. 關稅政策影響了台灣市場的外資動向, 特別是在大型金融機構的買超行為。
2. 美中貿易政策改變, 引發 PTT 中對於台灣晶片製造產業未來佈局的高度關注。
3. 科技人才的流動與招聘需求出現變化, 反映市場對未來產業發展的不確定性。
4. 國際政策(美日貿易關係)對台灣市場的影響深遠, 成為投資人關注的熱點。
5. 台灣股市的投資信心隨政策變動而波動, 反映在 PTT 討論熱度上。

## 四、結論

我們以 PTT 六個看板的討論為基礎，針對川普 2025 年關稅政策所引發的台灣社群輿論進行文本分析，發現這波政策所引起的反應已不僅是單一的經濟政策回應，而是牽動多層次社會關注的複合議題，涵蓋國際政治、供應鏈布局、科技發展與就業市場等領域。

透過 LDA、Guided LDA 與 BERTopic 三種主題模型，我們歸納出幾個重點現象：

1. AI 技術與美中科技對抗成為輿論核心主軸  
不論是 LDA 還是 BERTopic，皆顯示民眾對「AI 技術進展」與「美中科技競爭」高度關注，關鍵詞如「openai」「川普」「矽谷」「晶片」「關稅」等，呈現出台灣社群對於生成式 AI、半導體戰略與出口管制的集體焦慮與期待。
2. 半導體與高科技製造業的地緣壓力浮現  
討論熱點聚焦在台積電擴廠、政府補貼與南部科學園區的發展，同時關注與美日歐的合作，以及高雄等地的產業轉型，顯示台灣民眾對「製造業回流」下的台灣競爭力與生產鏈安全議題深具警覺。
3. 金融市場出現明顯波動感知與資金情緒轉移  
外資動向、法人買賣超、ETF 申購／贖回、金控公司（如元大、國泰、中信）皆成為重要語彙，顯示政策訊號直接牽動市場信心，並影響大眾對台股未來的預期。
4. 就業與人才流動議題隨政策升溫而浮出  
PTT 科技與工程師看板出現大量關於「遠端工作」、「外派」、「人才流動」、「簽證」與「裁員」等關鍵詞，反映出 AI 熱潮與關稅政策對工程師職涯、國際勞動市場及台灣留才／引才政策所造成的連鎖影響。
5. Guided LDA 分群更清晰，輿論面向更具焦點性  
相較於傳統 LDA 模型關鍵字重疊、主題邊界模糊，Guided LDA 的五個主題切分更為專一，能夠明確界定 AI 政策、美中對抗、股市資金動能、半導體供應鏈與就業市場等面向，顯示人為語意引導有助提升模型解釋力。
6. BERTopic 強化主題演變與語境遷移辨識  
BERTopic 展現不同模型下語意主軸的一致性，並突顯出「外資動向」、「國際供應鏈」、「技術人才」與「投資信心」四大脈絡，尤其顯示文本在時間跨度內語意聚焦與轉移的脈絡性。

我們觀察到，川普關稅政策不僅改變國際經濟格局，也重塑台灣輿論對「產業政策」、「就業信心」、「技術主權」與「經濟風險」的認知。這場討論已超越單一經濟議題，而是轉化為一場關於台灣未來角色定位與民間情緒集體轉向的社會溝通行動。