

野村投信 題目一

運用熱門主題及字組提升行銷轉換率之探索性研究

指導老師 蔡芸瑋 老師

指導業師 野村投信 張孝璿
野村投信 陳景堯

組員 蔣又中
邢恩展

目前做法：Top-Down

決定主推產品

野村優質基金

進行TA分析

台股基金投資人

選擇行銷管道

媒體新聞稿
網路廣告

規劃促銷內容

手續費優惠

反向做法：Bottom-Up

探知熱門投資話題
及關鍵字詞

話題分析

決定主推產品

規劃促銷內容

低利率

穩定收益產品

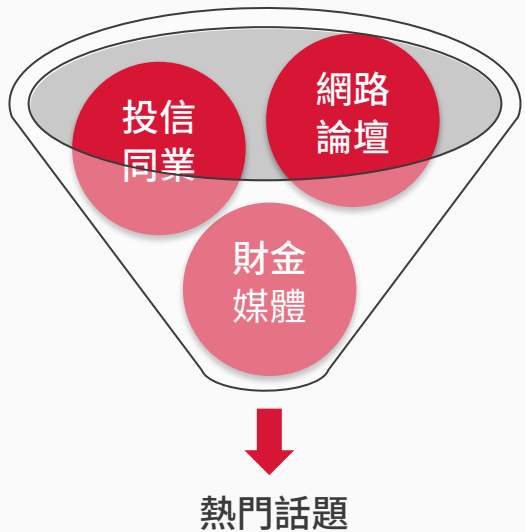
NN (L) 投資級公司債基金

手續費優惠

高收益債違約風險

穩定現金流

專案目標



透過蒐集網路論壇、新聞媒體以及投信同業網站內容，進行熱門字組探知與彙總，精準地掌握投資人間的熱門話題，使行銷上更加貼近消費者，提升行銷轉換率，進而改善marketing ROI。

資料來源



網路論壇

PTT基金版
強基金



財金媒體

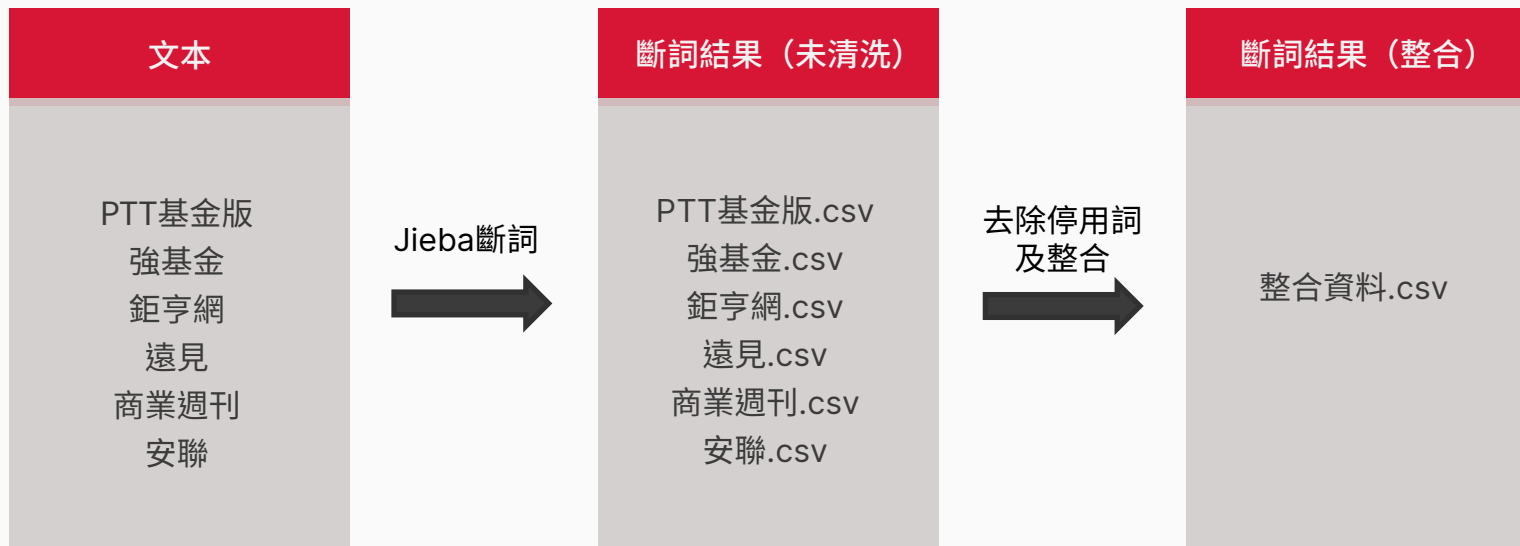
鉅亨網
遠見
今周刊
商業週刊



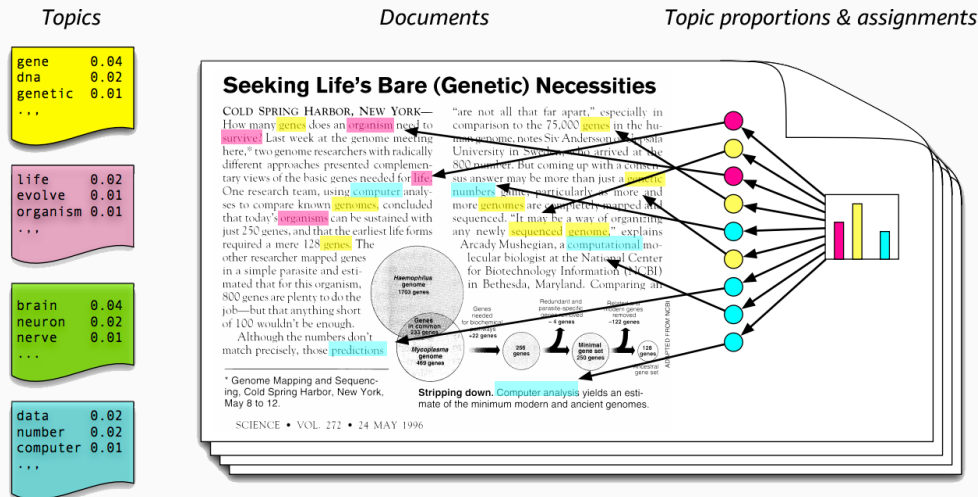
投信同業

安聯

資料來源

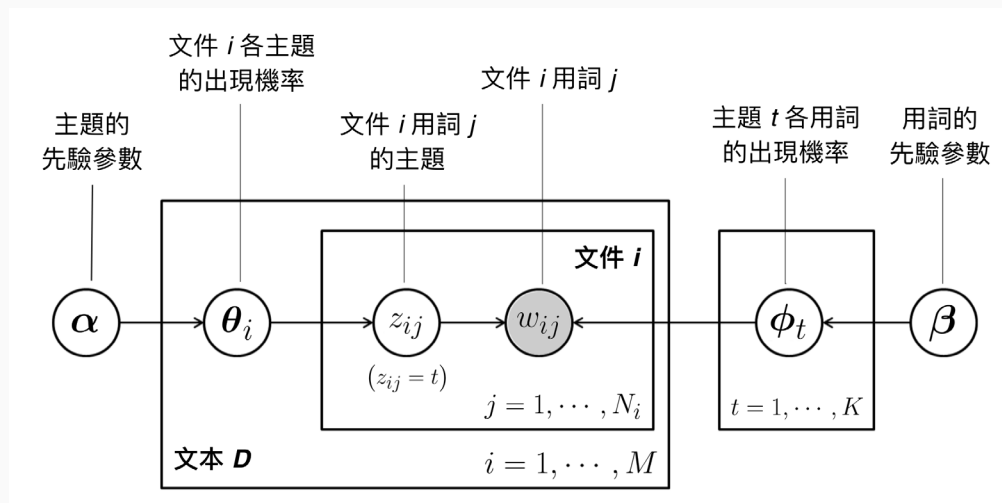


主題生成：Latent Dirichlet Allocation (LDA)



LDA是一種非監督式學習的文件主題模型，透過未觀察到的潛在變數來解釋為什麼數據的某些部分相似，或者可能隸屬於類似的「主題」群

主題生成：Latent Dirichlet Allocation (LDA)



- 每篇文件是由多個「主題 (Topic)」所組成
- 每個主題都可以由數個重要的「用詞 (Word)」描述，且相同的「用詞」可同時出現在不同主題之中

主題生成：pyLDAvis

- pyLDAvis 透過降維的方式將各主題的機率分布繪製在二維空間上
- 這能給我們兩樣重要的資訊：
 1. 是否產生有鑑別度的主題
 2. 主題彼此間的相關性。

主題生成：pyLDAvis

$$r(w, k|\lambda) = \lambda \log(\phi_{kw}) + (1 - \lambda) \log\left(\frac{\phi_{kw}}{p_{kw}}\right)$$

- 相關性分數有助於優先考慮更專屬於給定主題的字詞，使主題更加明顯
- lambda值越小時，第二項的權重越大，更專屬於該主題的字詞權重也越大
- 透過選定最佳lambda值，我們能更明確的定義主題

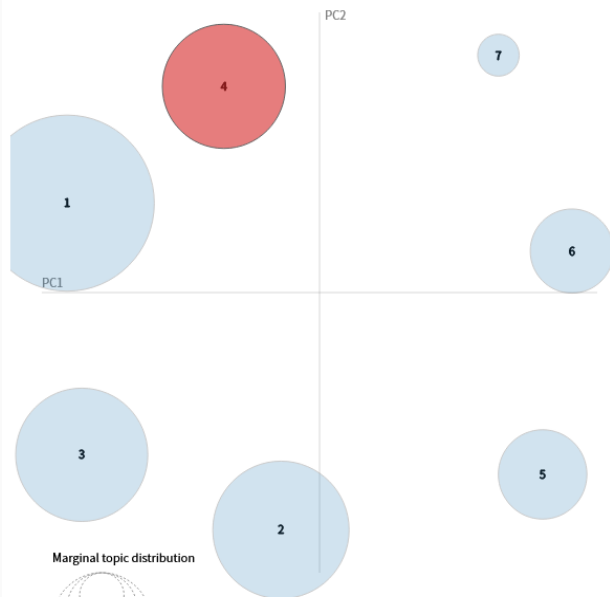
Selected Topic: 4 [Previous Topic](#) [Next Topic](#) [Clear Topic](#)

Slide to adjust relevance metric:⁽²⁾

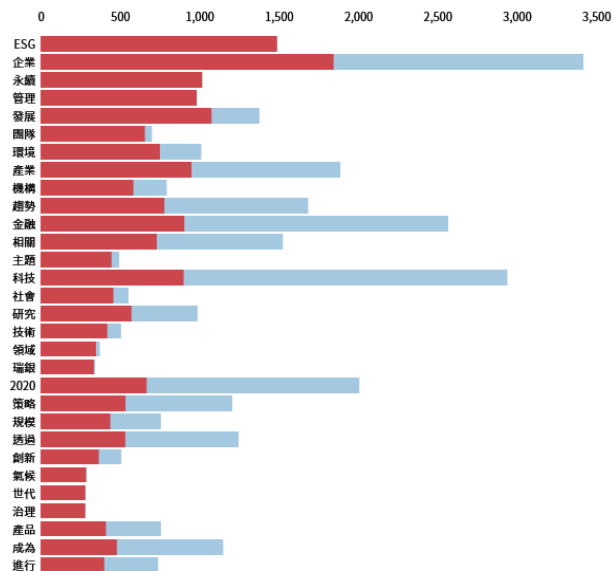
$\lambda = 0.6$

0.0 0.2 0.4 0.6 0.8 1.0

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 4 (15.4% of tokens)



Overall term frequency

Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)

2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

https://fintech0817206.s3.amazonaws.com/lda_06.html

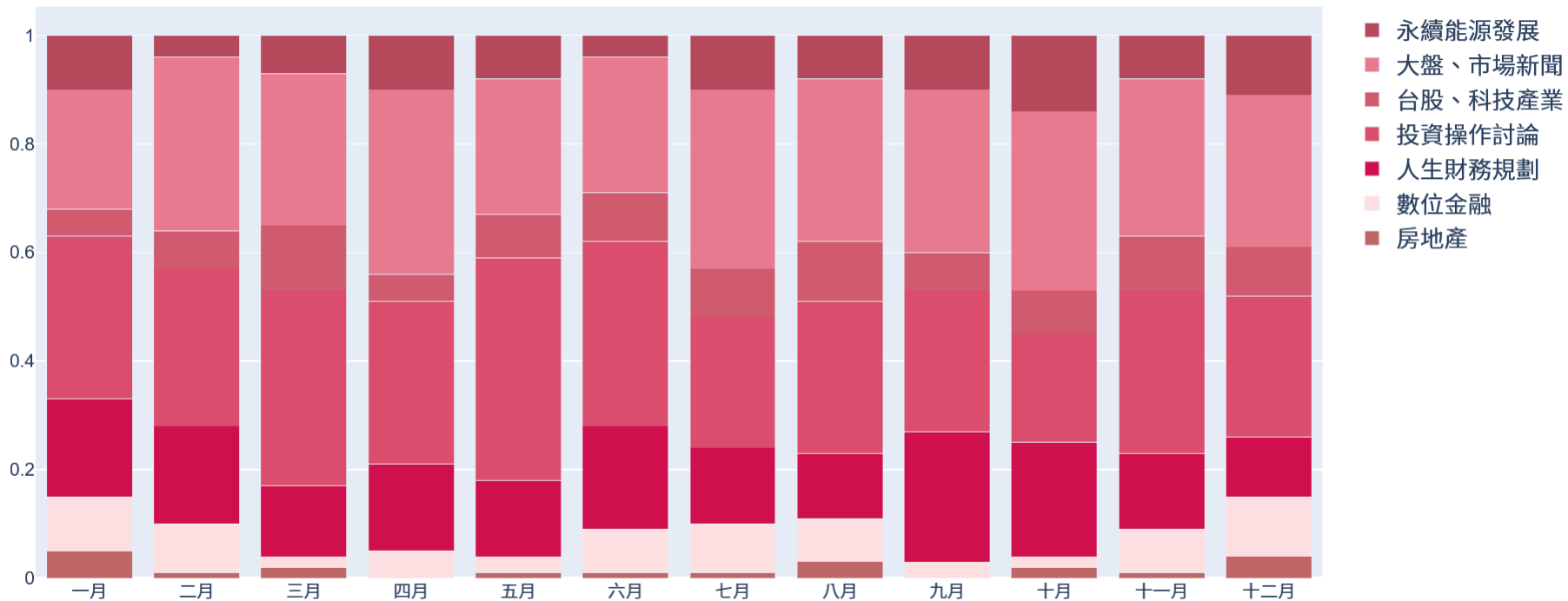


主題生成結果

當 $\lambda=0.6$

Topic 0	房地產	房價	房子	房屋	房市	台中	租金	不動產	台南	建商	貸款
Topic 1	數位金融	銀行	數位	客戶	服務	支付	金融	街口	信用卡	金管會	回饋
Topic 2	人生財務規劃	退休	理財	收入	帳戶	生活	需要	每個	方式	現金	選擇
Topic 3	投資操作討論	配息	定額	淨值	定期	南非幣	報酬率	單筆	績效	加碼	最近
Topic 4	台股、科技產業	台股	ETF	台積電	市值	直播	交易	指數	新台幣	上市	新高
Topic 5	大盤、市場新聞	股市	新興	上漲	政策	指數	資產	利率	預期	表現	下跌
Topic 6	永續能源發展	ESG	企業	永續	管理	發展	團隊	環境	產業	機構	趨勢

2020基金市場主題



結語

主要貢獻

- 透過文本分析找出消費者間的熱門主題與字組
- 利用文件主題模型將文章進行分類，方便行銷人員觀察話題熱度變化

改進方向

- 現階段礙於時間與運算資源，文本範圍為18個月左右，若能增加文本數量將能使主題更精確，分類準確度亦能再提升
- 將程式整合成網站形式，讓使用者不必熟練程式語言也能輕鬆使用

分工

蔣又中

爬蟲
資料清洗
LDA建置
分析結果繪製

邢恩展

主題發想
資料來源尋找
成果分析
簡報、影片製作