

假新聞判別

處理流程

資料來源 & 數據處理

- 資料：
[fake-and-real-news-dataset](#)
- 清理資料與合併

模型訓練

- Sciki-Learn
- Tensorflow
- Fasttext
- 測試準確度

新聞分類

- 比較三種機器學習
- 使用訓練好的模型
- 判別是否為假新聞

新聞摘要

- Gemini
- 摘要完辨別是否假新聞

SCIKIT-LEARN

資料來源 & 數據處理

- 資料 : [fake-and-real-news-dataset](#)
 - Fake.cvs: 23000篇假新聞
 - True.cvs: 21000篇真新聞
- 模型訓練只用text

	title	text	subject	date
1	title	text	subject	date
2	Donald Trump Sends Out Embarrassing New Year? Eye Messages	hi Donald Trump just couldn't wish all Americans a Happy New		31-Dec-17
3	Drunk Bragging Trump Staffer Started Russian Collusion Investigation	House Intelligence Committee Chairman Devin Nunes is g		31-Dec-17
4	Sheriff David Clarke Becomes An Internet Joke For Threatening To	ke it was revealed that former Milwaukee Sheriff David Clar	who was beir	has an email scandal of his t
5	Trump Is So Obsessed He Even Has Obama? Name Coded Into H	WOn Christmas day, Donald Trump announced that he woul		29-Dec-17
6	Pope Francis Just Called Out Donald Trump During His Christmas	ec Pope Francis used his annual Christmas Day message to re		25-Dec-17
7	Racist Alabama Cops Brutalize Black Boy While He Is In Handcuffs	gThe number of cases of cops brutalizing and killing people		25-Dec-17
8	Fresh Off The Golf Course, Trump Lashes Out At FBI Deputy Direc	r / Donald Trump spent a good portion of his day at his golf c		23-Dec-17
9	Trump Said Some INSANELY Racist Stuff Inside The Oval Office,	and In the wake of yet another court decision that derailed Don		23-Dec-17
10	Former CIA Director Slams Trump Over UN Bullying, Openly Suggest	ts Many people have raised the alarm regarding the fact that		22-Dec-17
11	WATCH: Brand-New Pro-Trump Ad Features So Much A** Kissing	t Just when you might have thought we d get a break from v		21-Dec-17
12	Papa John? Founder Retires, Figures Out Racism Is Bad For Busin	as A centerpiece of Donald Trump's campaign, and now his r		21-Dec-17
13	WATCH: Paul Ryan Just Told Us He Doesn't Care About Struggling	I Republicans are working overtime trying to sell their scam		21-Dec-17
14	Bad News For Trump ?? Mitch McConnell Says No To Repealing Ob	ama Republicans have had seven years to come up with a viable		21-Dec-17
15	WATCH: Lindsey Graham Trashes Media For Portraying Trump As	.The media has been talking all day about Trump and the R		20-Dec-17

6

流程圖

Scikit-learn TensorFlow Fasttext

Tokenizer

embedding

training process

test-train split

Shuffle and labeling datasets

Predicting result

15

LLM

- 目標：討論 LLM 生成新聞摘要（資料庫/網路新聞）後，是否仍能有效識別假新聞
- 方法：
 - 資料蒐集與處理：不同來源下的新聞（資料庫/網路新聞）生成摘要
 - 輸入摘要至三種分類模型（Scikit-learn、Fasttext、TensorFlow）判別真假新聞
 - 結果比較與分析

```
def generate_summary(title, content, max_chars=5000):  
    article_text = f"{title}\n\n{content}"  
    print(len(article_text))  
    article_text = article_text[:article_text.rfind('.') + 1] if len(article_text) > max_chars else article_text  
  
    system_prompt = (  
        "你是一位專業做摘要的人，請用英文做摘要（在五句話內），並保持專業和客觀性。"  
    )  
  
    user_prompt = f"Please summarize the following news article:\n\n{article_text}"  
    prompt = f"{system_prompt}\n\n{user_prompt}"  
  
    response = model_gemini.invoke(prompt)  
    return response.content.strip()
```

17

結果展示

<https://www.bbc.com/news/election-2017-40209087>



新聞標題：

Hung Parliament: Q&A guide to what happens when no-one wins the election

Gemini 生成摘要：

The 2017 UK general election resulted in a hung Parliament, leaving no party with a majority. Theresa May, as incumbent Prime Minister, intends to form a minority government, seeking support from the Democratic Unionist Party (DUP). A "confidence and supply" arrangement with the DUP is crucial for passing legislation and maintaining stability. Negotiations are underway, but the Conservative/DUP alliance faces potential challenges, and another election remains a possibility if the formation of a coalition is unlikely, and the uncertainty may complicate upcoming Brexit negotiations.

中文翻譯摘要：

2017 年英國大選產生了懸峙議會，沒有任何政黨取得過半數席次。
時任首相梅伊 (Theresa May) 打算組成少數政府，並尋求民主統一黨 (DUP) 的支持。
與民主統一黨達成「信任與支持」(confidence and supply) 的協議對於通過法案和維持政局穩定至關重要。
目前協商正在進行中，但保守黨/民主統一黨的聯盟面臨潛在的挑戰，如果政府變得無效，再次舉行選舉的可能性仍然存在。
組成聯合政府的可能性不高，而這種不確定性可能會使即將到來的英國脫歐談判變得更加複雜。

1/1 — 0s 33ms/step
({'__label__FAKE',), array([0.95349985])})

模型預測結果：

Scikit-Learn 預測：REAL (FAKE 機率：0.1018)
TensorFlow 預測：REAL (FAKE 機率：0.0000)
FastText 預測：FAKE (FAKE 機率：0.9535)

分析結果：

多數決結果：REAL

✓ 這篇新聞看起來是真新聞！

20

結果展示

<https://www.bbc.com/news/world-40215432>

BBC

UK election result: How the world reacted

10 June 2017

Share Save



新聞標題：

UK election result: How the world reacted

Gemini 生成摘要：

The UK election's unexpected outcome, where the Conservative party lost its majority,

中文翻譯摘要：

英國大選結果出乎意料，保守黨失去多數席位，在全球引發廣泛反應。歐盟領袖，包括容克和巴尼耶，都強調需要迅速

1/1 0s 134ms/step

模型預測結果：

Scikit-learn 預測：REAL (FAKE 機率：0.1557)

TensorFlow 預測：REAL (FAKE 機率：0.0002)

FastText 預測：REAL (FAKE 機率：0.0105)

分析結果：

多數決結果：REAL

✅ 這篇新聞看起來是真新聞！

21

心得與障礙排除分享

- LLM 中文摘要會大幅降低模型預測準確率
 - 可能原因：訓練資料以英文為主
- TENSORFLOW 模型表現最佳，正確率達 81.52%，深度學習模型可能能夠有效掌握新聞摘要的語意與真假訊息間的潛在關聯。
- FASTTEXT 在短文上的預測不如預期

22