

(一)摘要

得益於機器學習 (Machine learning) 的蓬勃發展，尤其是 google 提出的 self-attention 機制與 transformers 模型，attention 機制計算每個詞與詞之間的關聯，相較於傳統 RNN 方法更能理解語言中的一詞多義，transformers 則是實現這項技術的架構，為自然語言生成(Natural Language Generation)開啟了全新的里程碑。本研究將探索使用 BERT(Bidirectional Encoder Representations from Transformers)來生成網軍留言。藉由學習現有網路論壇上的網民對原始文章的留言，訓練出一個能夠學習網路用語，用字能夠與時俱進的留言生成模型。

(二)研究動機與研究問題

隨著社群網路發展、新媒體崛起，加快了資訊擴散速度，也是讓一般民眾不再只能被動接收訊息，同時也成為訊息的傳播者，不但可以透過社群平台分享自己對任何事物的看法，同時也能看到其他網路使用者的評論。原本是可以讓不同立場的使用者跨出同溫層互相交換意見，傾聽社會上不同意見的地方。

近年來許多網路行銷、公關公司，會使用大量假帳號以個人名義在 Facebook、PTT 等各大網路社群大量製造輿論、散布特定立場言論。利用民眾的從眾心理，影響一般使用者對事物的觀點與看法，藉此引導網路風向，將事件引導到對自己有利的局面。

有鑑於此，我們希望能開發出一套能夠生成網軍言論的模型，用以產生網軍留言。此項研究的動機在於提供能模擬網軍活動的環境。

以 PTT 上 100 篇關於韓國瑜的文章中推文標記為"噓"的推文為訓練資料進行測試。經過 20 次訓練後，分別就與韓國瑜相關文章與另一篇不相關的文章進行生成的結果 (圖 2.1, 2.2)

文章：[問卦] 有沒有高雄不捨韓國瑜，才不投他的卦？ 求有沒有中天新聞說 高雄人不投韓國瑜，是因為不捨得 他離開高雄的新聞影片？ 小弟找了好久都沒看到....	文章：送柯P腸胃藥、鮮花 王世堅嘆：史上最悲壯烙賽！ 史上最悲壯落賽，真的是好笑 不過就是 拉肚子 的小事，卻有人覺得是抗疫抗癌抗暴般的勇者 北市府大可好好利用這事件宣傳在抗疫以外 注重飲食衛生的重要性 不是只有戴口罩重要 可惜醫生市長第一時間只想到要復仇
-----	-----
網軍1：中天新聞不是叫假的[SEP] 網軍2：聞可以叫假的，但新聞影片質感真的有夠低[SEP] 網軍3：韓導的新聞也叫假的[SEP] 網軍4：韓市長是請假的新聞只看到有段沒錯的廢文[SEP] 網軍5：假新聞[SEP] 網軍6：韓市長是請假的新聞[SEP] 網軍7：假新聞[SEP] 網軍8：高雄人真的找得到？[SEP] 網軍9：新聞能信嗎[SEP] 網軍10：新聞能信嗎[SEP]	網軍1：醫生市長是副業[SEP] 網軍2：生是好笑的啊，醫生的任務就是讓健保們注重食[PAD]生意[SEP] 網軍3：醫生不要請假睡幾乎整天都不睡[SEP] 網軍4：醫生不要請假睡幾乎整天都不睡[SEP] 網軍5：睡菊還在睡[SEP] 網軍6：又是這個[SEP] 網軍7：醫生做的啊不心虛嗎[SEP] 網軍8：醫生做的都心虛嗎[SEP] 網軍9：醫生做的都心虛嗎[SEP] 網軍10：醫生做的都心虛嗎[SEP]

圖 2.1

圖 2.2

由實驗結果可知，主題與訓練資料接近的測試資料(圖 2.1)確實有能力產生跟文章相關且語句通順的負面評論。而(圖 2.2)雖然仍能從文章中抓到關鍵字"醫生"，但生成的內容還是以韓國瑜為主。

(三)文獻回顧與探討

過去在自然語言處理 (Natural Language Processing) 領域中。訓練架構主要是用 Recurrent Neural Networks 模型(圖 3.1)，RNN 缺點在於當字串過長時，相隔較遠的單字關聯性會遺失、還有梯度消失、難以平行運算等等問題。

由 google 於 2017 年提出的 self-attention 機制與 Transformer 架構[1, 2](圖 3.2)，大大改變 NLP 的訓練方式，self-attention layer 中所有 input vector 都會互相計算 attention，與 RNN 不同的是不需要依賴前一次的結果，每個 vector 計算 attention 時不會互相影響，因此能使用平行運算加速。

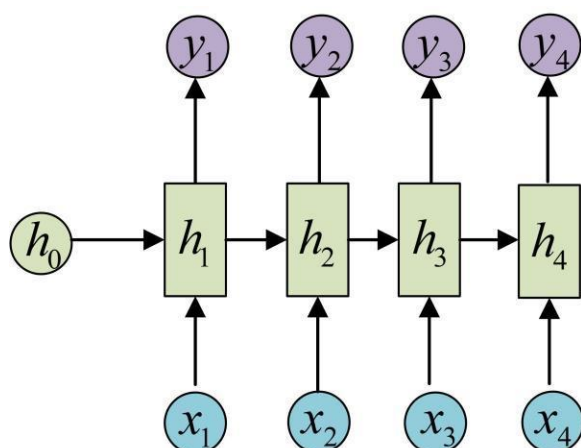
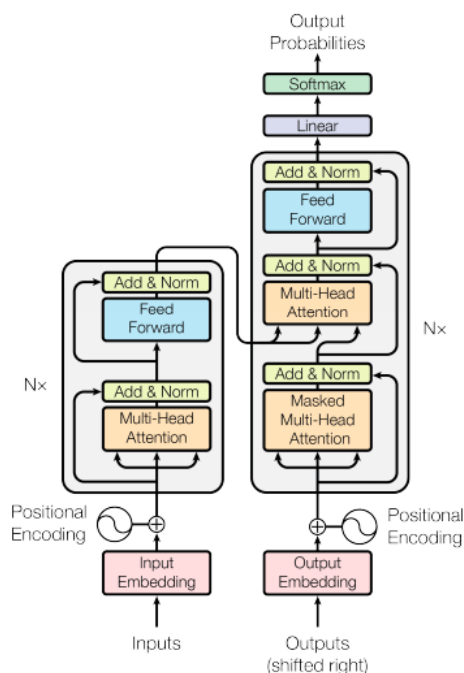


圖 3.1



The Transformer - model architecture.

圖 3.2: 論文[1]提出 2 的 transformers 架構

過去進行評論生成[3]是將任務拆分成 reading network 與 generate network(圖 3.3)，由閱讀網路抓出文章的關鍵字，再交由生成網路產生完整句子。得益於 transformers 強大的學習能力，能使用更一個簡單的網路就做到相近的成果，如果將任務拆分成更小的子任務，並且使用 transformers 把每部分都做到更好，預期能做到更符合真人評論的結果。

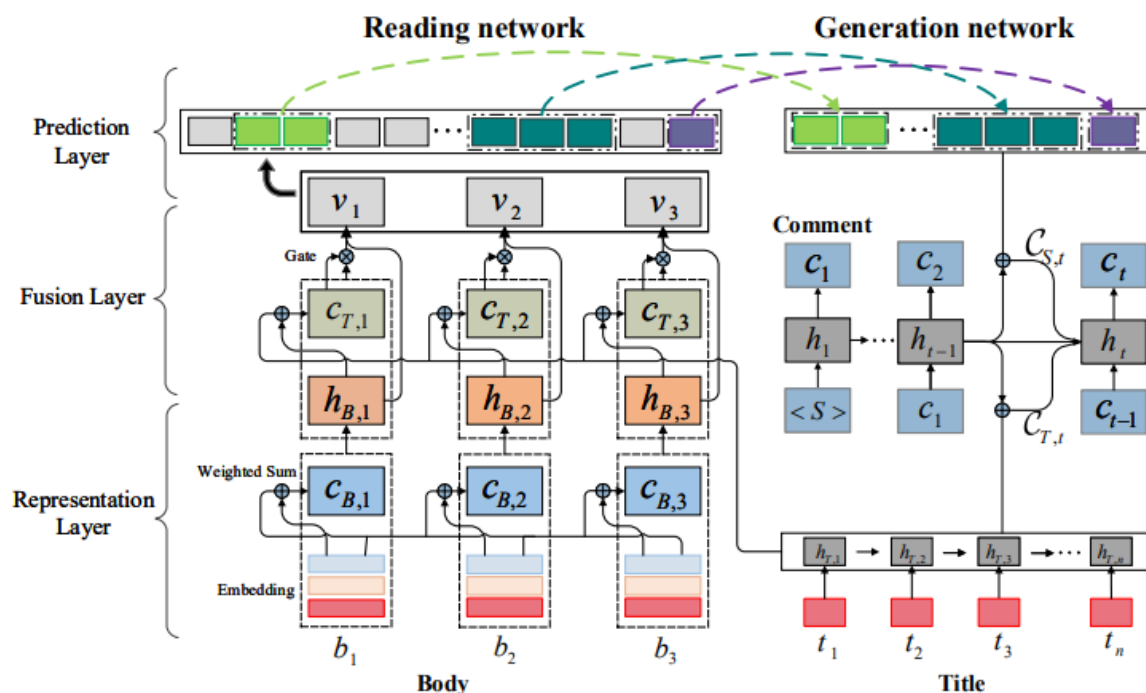


圖 3.3: 論文[4]使用的
兩種網路

(四)研究方法及步驟

我們選擇批踢踢實業坊(PTT)內做為訓練資料來源，使用 python beautifulsoup 套件蒐集看板文章、留言、留言標籤等資訊。並針對擷取的文章進行分類，抽取主題相近的文章、立場相似的留言，作為訓練時使用的資料。

```
{
  "article_id": "M.1578714664.A.718",
  "article_title": "Re: [問卦] 不帶印章選務人員強制蓋手印有違規嗎",
  "author": "strike519 (噓我者 宅友也)",
  "board": "Gossiping",
  "content": ": 標題: [問卦] 不帶印章選務人員強制蓋手印有違規嗎 : 時間: Sat Jan 11 10:16:45 2020 :",
  "date": "Sat Jan 11 11:51:02 2020",
  "ip": "59.126.22.29",
  "message_count": {
    "all": 19,
    "boo": 16,
    "count": -14,
    "neutral": 1,
    "push": 2
  },
  "messages": [
    {
      "push_content": "你能不能刪一些...",
      "push_ipdatetime": "180.217.83.181 01/11 11:51",
      "push_tag": "噓",
      "push_userid": "DON3000"
    },
    {
      "push_content": "你可以推文",
      "push_ipdatetime": "61.228.43.47 01/11 11:52",
      "push_tag": "噓",
      "push_userid": "aggressorX"
    },
    {
      "push_content": "多菜",
      "push_ipdatetime": "61.230.22.199 01/11 11:52",

```

圖 4.1

合併文章內文(content)跟留言(push_content)組合成一個新的字串，經過 tokenizer 把字串 encode 成 BERT 字表上相應的 ids 作為 BERT 的輸入(圖 4.2)。

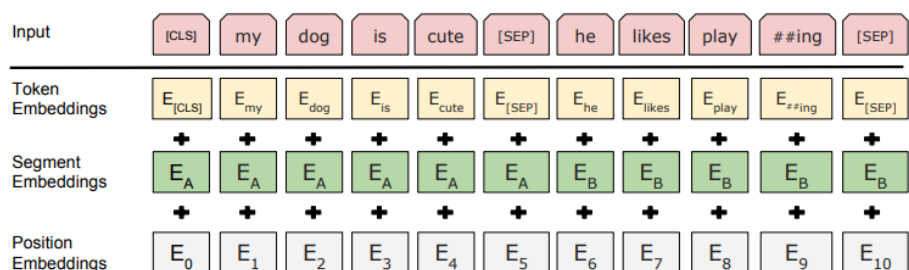


Figure 2: BERT input representation. The input embeddings are the sum of the token embeddings, the segmentation embeddings and the position embeddings.

圖 4.2: BERT 的 input 形式

利用 Facebook 所開發的神經網路框架 Pytorch，使用 pytorch 版本 bert-base-multilingual-cased 預訓練模型，預計使用兩種不同方式進行實驗。目前遇到的問題，主要是無法對所有文章產生與之對應的回覆和句子結構不通順，前者於訓練時加入其他種類的文章可以改善，後者則是預計花更多時間訓練，產生句子後先判斷語句是否通順完整，剔除雜亂的評論等等後處理。

第一種方式為使用克漏字填空的方式讓模型逐字預測評論種每個字。在文章最後加入特殊符號"[MASK]"經過 transformers 得到 embedding 後使用 classifier model 預測出機率最高的字"**r**"，把"[MASK]"帶換成"**r** [MASK]"，重複讓 model 預測直到出現代表句子結尾的特殊符號"[SEP]"出現 (圖 4.3)。

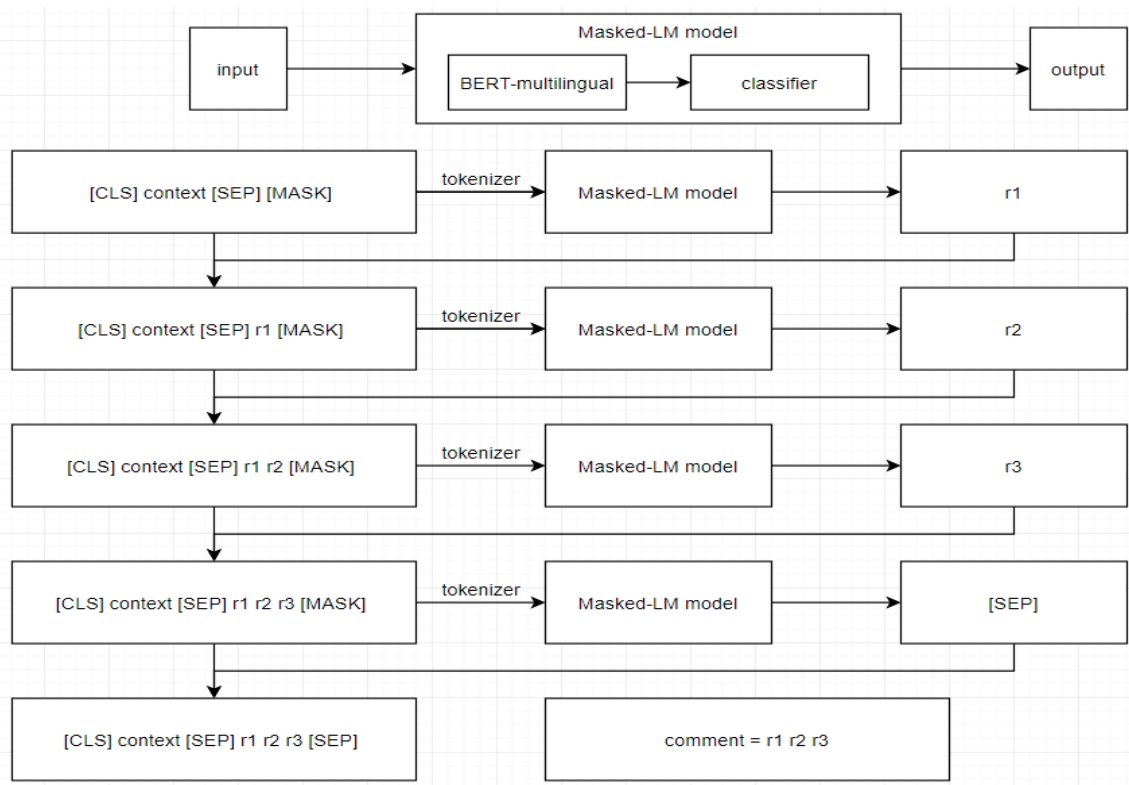


圖 4.3

第二種為將原任務分成拆成"閱讀文章找出關鍵字"與"使用關鍵字生成評論"的兩個子任務分別交由兩個 transformers 完成。先由 keyword model 找出文章的關鍵字再由 generate model 生成評論(圖 4.4)

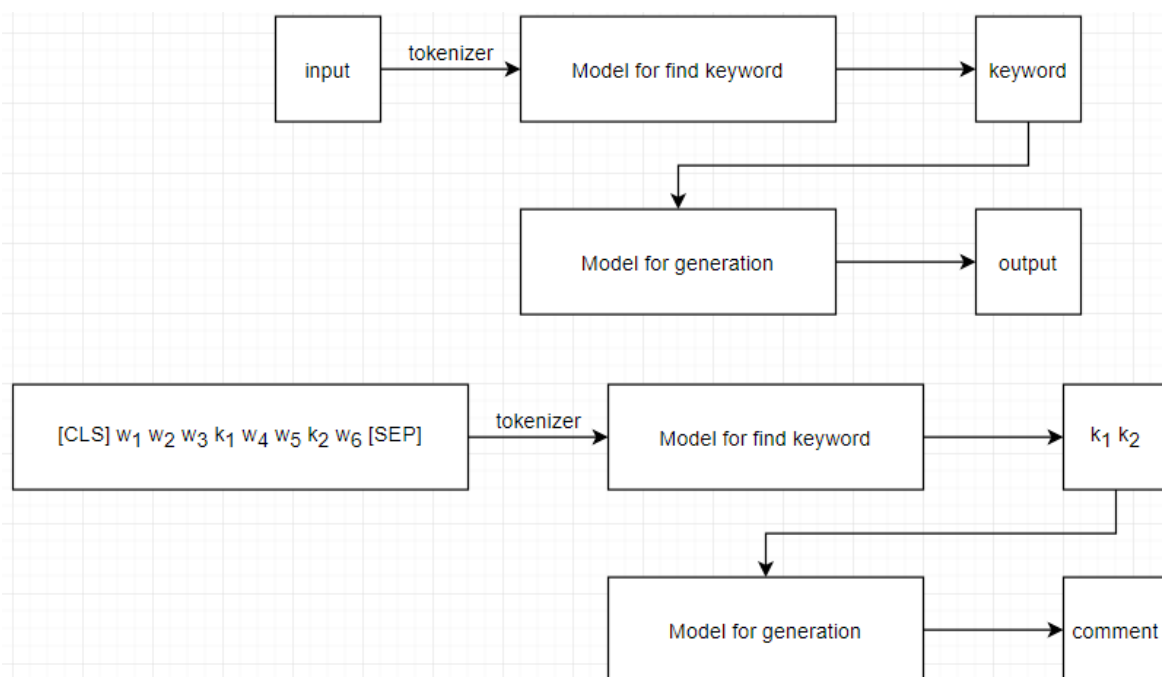


圖 4.4

(五)預期結果

本研究預期在給定任一真實文章後，能夠產生與其陣營對立的網軍評論(圖 5.1)。未來期望此系統能夠模擬網軍活動時的環境，提供偵測真實網軍時所需的資料，以協助執法機關判斷網軍是否有對議題進行干涉。

文章：[爆卦] 韓國瑜開漫威記者會啦！！
 高雄市長韓國瑜昨(13)日爆料表示，知名美國漫威公司有意將他奮鬥的過程變成漫畫，但漫威官方卻回覆完全不知情，引發爭議。對此，韓國瑜將於下午親自說明現任AAE執行長 David Martin計畫將市長的故事以漫威風格製成作品等事宜。

網軍1：這個需要開記者會 ?????
 網軍2：吃飽太閒 整天幹一堆跟高雄無關的鳥事
 網軍3：就不是漫威啊 韓國瑜每天吹牛啊
 網軍4：炒新聞 反正有鐵粉護航

圖 5.1

Input	predict
[CLS] context [SEP] [MASK]	這
[CLS] context [SEP] 這 [MASK]	個
[CLS] context [SEP] 這 個 [MASK]	需
[CLS] context [SEP] 這 個 需 [MASK]	要
[CLS] context [SEP] 這 個 需 要 [MASK]	開
[CLS] context [SEP] 這 個 需 要 開 [MASK]	記
[CLS] context [SEP] 這 個 需 要 開 記 [MASK]	者
[CLS] context [SEP] 這 個 需 要 開 記 者 [MASK]	會
[CLS] context [SEP] 這 個 需 要 開 記 者 會 [MASK]	?
[CLS] context [SEP] 這 個 需 要 開 記 者 會 ? [MASK]	?
[CLS] context [SEP] 這 個 需 要 開 記 者 會 ?? [MASK]	?
[CLS] context [SEP] 這 個 需 要 開 記 者 會 ??? [MASK]	[SEP]
[CLS] context [SEP] 這 個 需 要 開 記 者 會 ??? [SEP]	
[CLS] context [SEP] 評論 1 [SEP] [MASK]	吃
[CLS] context [SEP] 評論 1 [SEP] 吃 [MASK]	飽
[CLS] context [SEP] 評論 1 [SEP] 吃 飽 [MASK]	太
[CLS] context [SEP] 評論 1 [SEP] 吃 飽 太 [MASK]	閒
.	.
.	.
.	.

(六)參考文獻

- [1] Attention Is All You Need Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhi, 2017
- [2] BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, 2018
- [3] ptt-web-crawler (PTT 網路版爬蟲) <https://github.com/jwlin/ptt-web-crawler>
- [4] Read, Attend and Comment: A Deep Architecture for Automatic News Comment Generation Ze Yang, Can Xu, Wei Wu, Zhoujun Li, 2019

(七)需要指導教授指導內容

定期與老師討論研究進度與相關機器學習之背景知識，並用實驗室之 GPU 伺服器以訓練神經網路模型