10994015 資管碩二 李承諺

心得:

本次心得我撰寫許政堯同學的**深度學習結合類別不平衡處理於檢測虛假評論**

，評論對我來說是非常重要的，隨著行動網路的普及化，我在用餐前或去某個景點之前都會上網閱讀該地的評論，有時結束友會對此地做出評論，而對我而言相較於撰寫評論，我更多的時間是在查詢評論，但網路世界中資料不平衡(虛假資料)越來越多，不管是評論，甚至連文章、影片、圖片都隨著科技的發展陸續出現偽造的資料，要想正確抓取這些偽造的資料會越來越難，因為科技的發展作假的能力越來越強，例如近年還火紅的GAN(生成對抗網路)就可以偽造出許多假資料如假評論、假圖片、假文章來以假亂真，人類也逐漸分不清哪些到底是人類還是電腦產生的，例如GAN火紅的生成人臉的技術，說真的電腦生成的人臉與真實人臉根本一模一樣，無法分辨，所以我很期待政堯同學未來的研究，若能對這塊領域的研究有做出貢獻的話，我認為一定是很值得期待的一篇論文。

1. 基本資料

駱亞聖（2021）。深度學習應用於生成客製化餐廳評論之研究。國立檯北科技大學電機工程系碩士班

1. 研究問題

隨著科技日新月異及網路的普及，以吃飯為例，在前往某家餐廳用餐前往往都會先看評論，但大多數人都不想花費過多的時間在評論上，卻有想得知其他人的評論，因此此論文以餐廳評論作為研究，提出一個簡易的評論生成系統，分析各店家餐膯的評論等等，透過簡單且能夠符合顧客想法的客製化評論讓大家在用餐後能夠快速地對餐廳做出評論心得，以幫助後人做為參考評論的依據，而此論文使用兩種非常火紅的深度學習的模型，一種是Transformer模型一種是BERT模型，並在GoogleMaps上蒐集餐廳評論資料進行訓練，來觀察哪種模型對此研究訓練得更優秀。

1. 研究方法/提出的方法論

此研究方法分為四種不同階段，分別為資料蒐集、文字預處理、Transformer模型訓練、BERT模型的訓練。

資料蒐集擷取GoogleMpas在台灣地區的9000多家餐廳共150萬筆左右的評論，先後剔除掉少於10個字過於簡短的評論以及大於256個字過於冗長極端的評論，接著從餐廳評分中的5個星級評論中每個挑選四萬筆共二十萬筆評論作為訓練資料。

接著進行文字預處理，一般做NLP的時候通常會將文字進行量化，也就是將文字轉化為數值，而此研究會根據文本的字詞量建立token字典，每個字詞都會有對應的token號碼。最後將處理好的資料分別丟入Transformer模型以及BERT模型進行文字的生成，並以生成評論的多樣性、精確度、相關度來做為評判模型優劣的標準。

1. 實驗結果:

實驗結果分別觀看兩種深度學習模型進行比較。

1. Transformer:

流暢度:

在Transformer評論生成中，大部分評論的流暢度都很好，但還是有少數語句並不是非常通順，簡單來說就是有語意邏輯上或矛盾的問題存在，比如說有一句話是:”*菜單上的肉品都很棒，服務態度也很好，但是肉 質真的不好，不是很好吃，而且加價的肉品有點少，不過整體來說是很不錯的。*”，看得出第一句菜單上的肉品都很棒與第三句的肉質真的不好產生了語意矛盾的問題，還有一則評論:” 店家態度很差，店裡的客人也是一樣，我們一直 以為我們有錯，結果店員居然在我們後面來的 客人都進來了，店員還說我們沒訂位，為什麼不 是讓我們先到的，這樣對待其他客人嗎？我們 還是訂位的？”，可以看出整段話的邏輯非常有問題，推測可能是在訓練過程模型讓某幾個字過於頻繁出現所導致的狀況。

準確度:

評論的準確度是以評論為正面還是負面所分類的，以評論3星級為分界點，4、5星為正向評論，1、2星為負面評論，而Transformer在評分4的時候會有點模糊不清，會先以正面評論作為開頭，之後又會有與其轉折的地方，對正負面的定義就較為模糊，例如一則四星評論:” *生意很好，東西也好吃。但是店家態度很差，明明就沒有客滿，還說等一下。 等很久，後面來的客人都吃飽離開了，才慢慢的進來，感覺很差！*”，前面為正面但後面轉折為負面卻還是四星評論，也認為此評論的分類較為模糊，從實驗結果來看Transformer對評分的準確度是較為不拿手的。

多樣性:

生成評論的多樣性為評比評論好壞的重要指標，主要原因就是每種評論都要有變化性，不管是時間、地區、不同人、不同餐廳、不同口氣都應該有不同變化的評論，由五則評論來觀察:

1. *服務態度超差，點餐時也沒介紹，點到一半就在趕人，服務生也沒有介紹菜單，整個 用餐心情超差，完全不推。*
2. *服務態度好，餐點好吃。但有位女服務生態度不是很好，不過也許因為有帶長輩或是 小朋友，餐點都還滿意。*
3. *今天吃了三種口味的 xo 醬炒麵，都很好吃！份量也夠，燒賣跟蝦餃都很推薦！*
4. *菜單上寫著假日營業到 22: 00 但是 20: 30 就不收客了感覺不是很好。*
5. *海鮮很新鮮，菜色也不錯，價格也很公道，是一家會想再來的海鮮餐廳。*

可以從以上五則評論看到以Transformer生成的評論還是具有一定的多樣性，對此我們判定Transformer在評論的多樣姓的表現是還不錯的。

1. BERT:

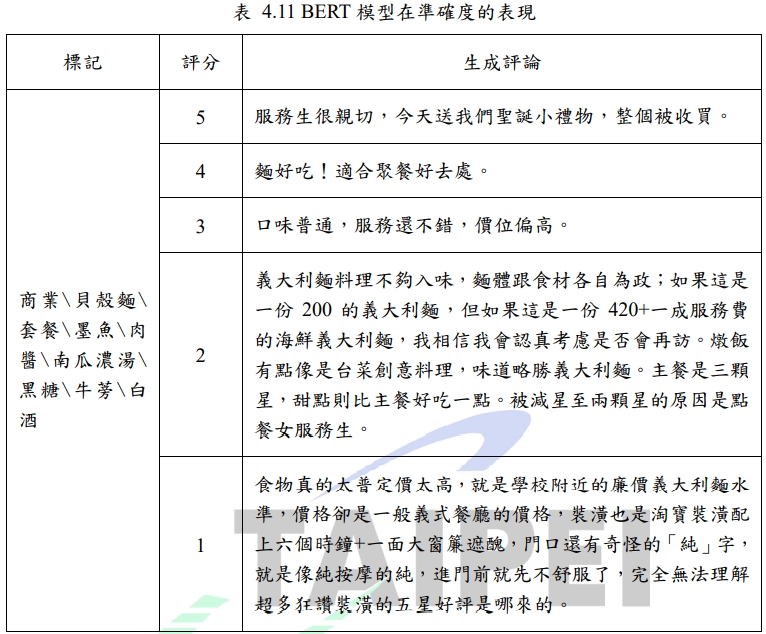
相關度:

一般評論都需與餐廳，也就是我們輸入的標記有關連性，以達到客製化評論的需求，而BERT就可以做到這點，從BERT的實驗結果來看，BERT所生成的模型都具有一定的關聯性，在關聯性這塊表現是相當成功的。

流暢度、準確度:

在BERT的流暢度方面，我們認為已接近人類的表現，前後矛盾文以及邏輯不通的情況下幾乎沒有發生，猜測可能為BERT預訓練模型有經過大量的非監督訓練，讓BERT的生成評論都達到很好的結果。

其中一則:” *永和牛肉麵老店，只有賣牛肉麵牛肉湯跟簡單的小菜， 麵條有家常麵寬麵細麵冬粉多種選擇，口味不錯，喜 歡吃牛肉麵的人值得一試，不過生意非常好，用餐時 間可能需要排隊。牛肉家常麵(小) $ 120。*”。準確度來說也幾乎能夠掌握，每個星級的評論幾乎都符合我們的期望。下圖為參考資料。



多樣性:

由以下五則評論來分析多樣性:

1. *服務態度極差，主食不熱沒味道，連茶碗蒸裡面還有頭髮，這樣的價位不應該是這樣 的服務，咖哩非常不好吃，炸排骨都是油，咖哩非常稀，口感極差，這樣的服務非常 不推薦。*
2. *母親節訂位 7: 30，到現場去超傻眼，一堆更之前訂位的人都還沒入席，亂成一片。 帶自己的媽媽來吃個大餐，在外頭等超過 40 分鐘都還沒入席超扯。*
3. *好友的婚禮在這裡舉辦，真的很正點。*
4. *家長可以坐在餐桌就能看到小孩很方便！*
5. *很少給差評，偶爾點有特色的餅皮，但絕對不是地雷，很像是在吃路邊攤的麵包加熱 出來賣的。*

由以上五則平輪來觀察，可以發現多樣性十分足夠，很直覺得會認為這五篇都是由不同人所發表出來的評論，實際上全都是由AI生成的，因此對此我們認為BERT的多樣性評論效果是非常好的。

結論:

在實驗中可以看到Transformer大致可以掌握大方向，但許多細節處理的並不是很好，相較於BERT細節就能夠處理得非常好，在多樣性表現也較突出，實驗結果顯示 BERT 模型在生成客製化餐廳評論上的表現優於 Transformer 模型。從實驗也見證了深度學習模型生成語言已經能夠非常成熟，也間接證實了預訓練語言處理慢慢走向主流的原因。