

DLHLP HW 4 - 1

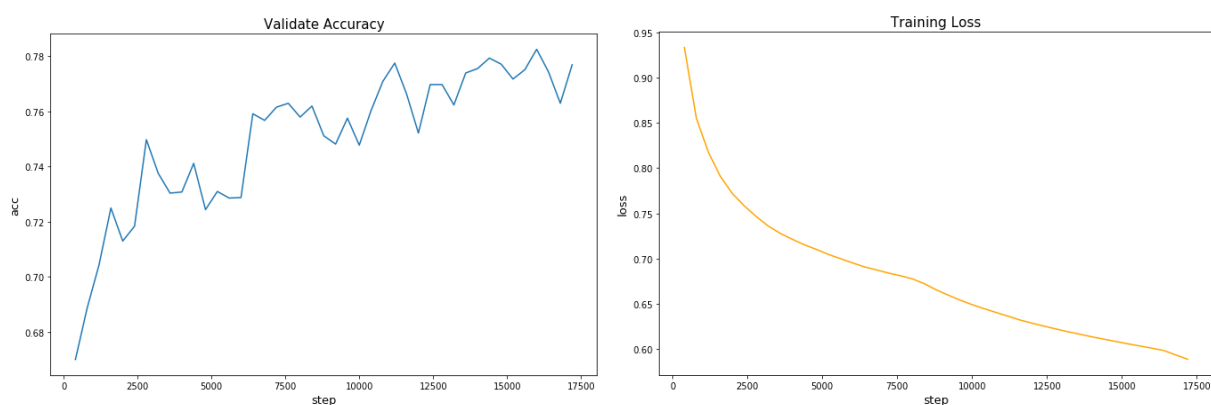
組別：wubinary

學號：r08921040, r08942087

姓名：徐均筑, 吳彬睿

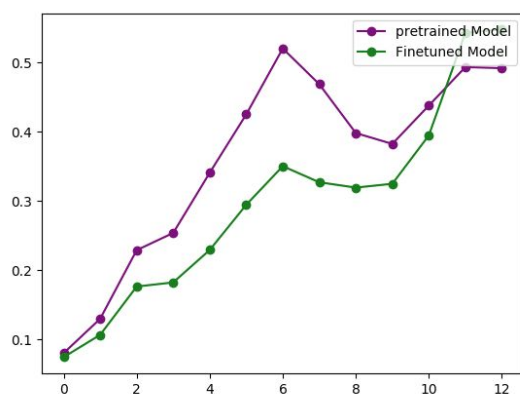
Part 1:

1.(1%) 請將training 的 loss curve 跟 test accuracy 截圖 放在下面，並簡單交待一下你的training 過程

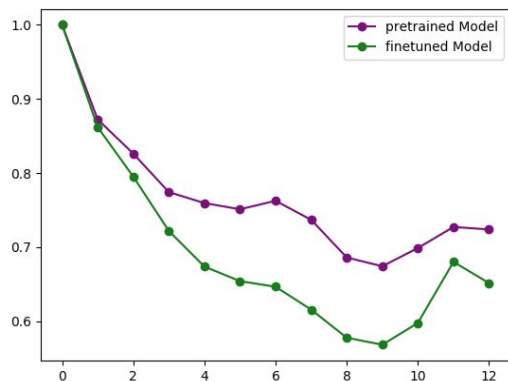


Fine-tune BERT 只需要少少的幾個 epoch 就足夠，且 learning rate 不能太大。我們將 Batch size 設為 48， learning rate 設為 $1e-5$ ，並使用大約 500 個 warm-up steps，訓練 3 個 epoch 便停止。

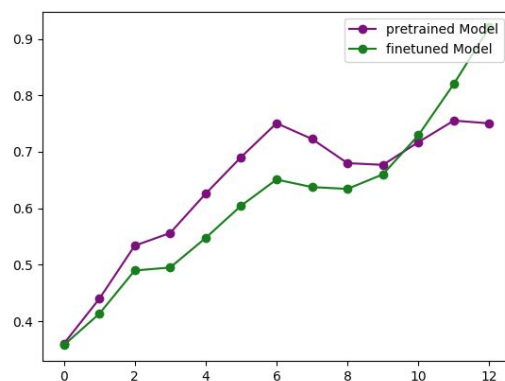
2.(1%) 請根據 Anisotropy 的定義 以及助教提供的範例圖示，畫出 0 - 12 層各層的 Anisotropy 數值 所連起來的線， pretrained model的版本請用紫色線畫出來， finetune model的版本請用綠色線畫出來



3.(1%) 請根據 self-similarity 跟 intra-sentence similarity 的定義 以及 助教提供的範例圖示，畫出兩張圖片，一張是 0 - 12 層 各層的 self-similarity 數值所連起來的線，一張是 0 - 12 層各層的 intra-sentence similarity 數值所連起來的線，每張圖片，都會有兩條線，一條線是 pretrained model 的版本，另一條線是 finetune model 的版本，pretrained model 的版本請用紫色線畫出來，finetune model 的版本請用綠色線畫出來

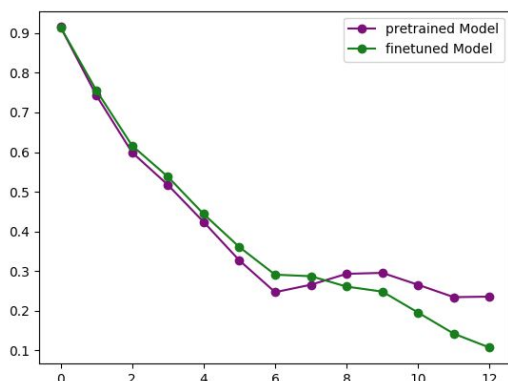


(self-similarity)

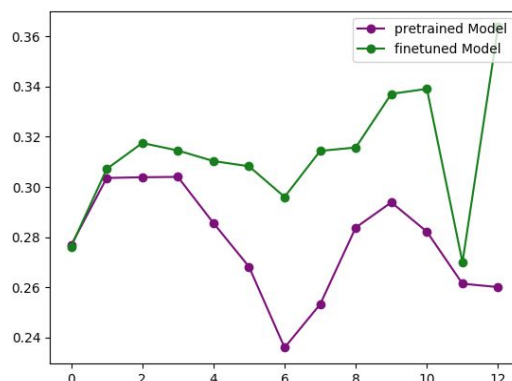


(intra-sentence similarity)

4.(2%) 請把 self-similarity 以及 intra-sentence similarity 有減去各層的 anisotropy 的圖(adjust version)畫在本題(配色如同第2題並標示清楚)，並且比較一下兩張圖和前一題做出來的兩張圖的差異，試著解釋一下 finetune 前後的變化。



(self-similarity)



(intra-sentence similarity)

- Observation

1. Higher layers 中的 representations 通常比較 context-specific , 而 contextualization 又伴隨著 high anisotropy , 因此在若未 adjust for anisotropy , 在第10 - 12 層的 self-similarity 相較前幾層高。
2. 不似 anisotropy , contextualization 通常不一定會伴隨著 high intra-similarity , 因此 adjust version 中 intra-similarity 並沒有一定的趨勢。
3. 相較於 pretrained model , finetuned model 的 self-similarity 較低、intra-similarity 較高 , 顯示 finetuned model 的 representations 比較 context-specific 。

Part 2:

1. (3%) Segment these sentences:

● Result

1. 我,一直,親自,指揮,、,親自,部署,, ,我,相信,只要,我們,堅定,信心,、,同舟共濟,、,科學,防治,、,精準,施策,, ,我們,一定,會,戰勝,這,一,次,疫情,。
2. 這,個,聲明,讓,我,再次,想起,了,安徒生,的,童話,《,皇帝,的,新裝,》,。
3. 希望,他們,能夠,聽,一,聽,這,個,忠告,, ,不,要,再,信口雌黃,地,抹黑,, ,居心叵測,地,挑撥,, ,煞有介事,地,恫嚇,。
4. 有關,部門,當然,就,是,有關,的,部門,了,。 ,無關,的,就,不,能,稱為,有關,部門,, ,所以,我,建議,你,還,是,要,向,他們,詢問,。
5. 不,要,搞,奇奇怪怪,的,建築,。
6. 現在,提請,表決,。 ,同意,的,代表,請,舉手,。 ,請,放下,; ,不,同意,的,請,舉手,。 ,沒有,; ,棄權,的,請,舉手,。 ,沒有,。 ,通過,!
7. 輕關易道,, ,通商,寬衣,。
8. 人均,國內,生產,總值,接近,八千萬,美元,。
9. 我,青年,時代,就,對,法國,文化,抱有,濃厚,興趣,, ,法國,的,歷史,、,哲學,、,文學,、,藝術,深深,吸引,著,我,。 ,讀,法國,近現代史,特別是,法國,大,革命史,的,書籍,, ,讓,我,豐富,了,對,人類,社會,政治,演進,規律,的,思考,。 ,讀,孟德斯鳩,、,伏爾泰,、,盧梭,、,狄德羅,、,聖西門,、,傅立葉,、,薩特,等,人,的,著作,, ,讓,我,加深,了,對,思想,進步,對,人類,社會,進步,作用,的,認識,。 ,讀,蒙田,、,拉封丹,、,莫里哀,、,司湯達,、,巴爾扎克,、,雨果,、,大仲馬,、,喬治·桑,、,福樓拜,、,小仲馬,、,莫泊桑,、,羅曼·羅蘭,等,人,的,著作,, ,讓,我,增加,了,對,人類,生活,中,悲歡離合,的,感觸,。 ,冉阿讓,、,卡西莫多,、,羊脂球,等,藝術,形象,至今,仍,栩栩如生,地,存在,於,我,的,腦海,之中,。 ,欣賞,米勒,、,馬奈,、,德加,、,塞尚,、,莫內,、,羅丹,等,人,的,藝術,作品,, ,以及,趙無極,中,西,合璧,的,畫作,, ,讓,我,提升,了,自己,的,藝術,鑑賞,能力,。 ,還有,, ,讀,凡爾納,的,科幻,小說,, ,讓,我,的,頭腦,充滿,了,無盡,的,想像,。
10. 因為,我,那,時候,, ,扛,兩百,斤,麥子,, ,十里山路,不,換肩,的,。

2. (2%) 從助教給的例子中，我們會發現機器遇到標點符號時必定預測其label為S。如果去除標點符號，是否會對句子的segmentation造成影響呢？請對上述10句的無標點符號版本進行segmentation，並敘述你的觀察。

- Result

1. 我,一直,親自,指揮,親自,部署,我,相信,只要,我們,堅定,信心,同舟共濟,科學,防治,精準,施策,我們,一定,會,戰勝,這,一,次,疫情
2. 這,個,聲明,讓,我,再次,想起,了,安徒生,的,童話,皇帝,的,新裝
3. 希望,他們,能夠,聽,一,聽,這,個,忠告,不,要,再,信口雌黃,地,抹黑,居心叵測,地,挑撥,煞有介事,地,恫嚇
4. 有關,部門,當然,就,是,有關,的,部門,了,無關,的,就,不,能,稱為,有關,部門,所以,我,建議,你,還,是,要,向,他們,詢問
5. 不,要,搞,奇奇怪怪,的,建築
6. 現在,提請,表決,同意,的,代表,請,舉手,請,放下,不,同意,的,請,舉手,沒有,棄權,的,請,舉手,沒有,通過
7. 輕關,易道,通商,寬衣
8. 人均,國內,生產,總值,接近,八千萬,美元
9. 我,青年,時代,就,對,法國,文化,抱有,濃厚,興趣,法國,的,歷史,哲學,文學,藝術,深深,吸引,著,我,讀,法國,近現代史,特別是,法國,大,革命史,的,書籍,讓,我,豐富,了,對,人類,社會,政治,演進,規律,的,思考,讀,孟德斯鳩,伏爾泰,盧梭,狄德羅,聖西門,傅立葉,薩特,等,人,的,著作,讓,我,加深,了,對,思想,進步,對,人類,社會,進步,作用,的,認識,讀,蒙田,拉封,丹莫里哀,司湯達,巴爾扎克,雨果,大仲馬,喬治·桑,福樓拜,小,仲馬,莫泊桑,羅曼·羅蘭,等,人,的,著作,讓,我,增加,了,對,人類,生活,中,悲歡離合,的,感觸,冉·阿讓,卡西莫多,羊脂球,等,藝術,形象,至今,仍,栩栩如生,地,存在,於,我,的,腦海,之中,欣賞,米勒,馬奈,德加,塞尚,莫內,羅丹,等,人,的,藝術,作品,以及,趙無極,中,西,合璧,的,畫作,讓,我,提升,了,自己,的,藝術,鑑賞,能力,還有,讀,凡爾納,的,科幻,小說,讓,我,的,頭腦,充滿,了,無盡,的,想像
10. 因為,我,那,時候,扛,兩百,斤,麥子,十,里,山路,不,換肩,的

- Observation

從結果而言無太大影響，model 能夠由文字內容來判斷何時應該斷句，而非單純依照標點符號進行斷句。