

# DLHLP HW4-3 Report

組長 Github ID: openopentw

組員 (姓名 + 學號): 鄭淵仁 R08922067

- 1 Please give some examples predicted correctly and incorrectly respectively. At least one for each case is required. Screenshot recommended.

## 1.(a) Correctly

```
In [4]: dev['data'][0]['paragraphs'][1]
Out[4]:
{'context': '梵語在社交中口頭使用，並且在早期古典梵語文獻的發展中維持口頭傳統。在印度，書寫形式是當梵語發展成俗語之後才出現的；在書寫梵語的時候，書寫系統的選擇受抄寫者所處地域的影響。同樣的，所有南亞的主要書寫系統事實上都用於梵語文稿的抄寫。自19世紀晚期，天城文被定為梵語的標準書寫系統，十分可能的原因是歐洲人有用這種文字印刷梵語文本的習慣。最早的已知梵語碑刻可確定為公元前一世紀。它們採用了最初用於俗語而非梵語的婆羅米文。第一個書寫梵語的證據，出現在晚於它的俗語的書寫證據之後的幾個世紀，這被描述為一種悖論。在梵語被書寫下來的時候，它首先用於行政、文學或科學類的文本。宗教文本口頭傳承，在相當晚的時候才「不情願」地被書寫下來。',
'id': '1147-6',
'qas': [{'id': '1147-6-1',
'question': '印度的梵語的書寫形式是到何時才出現的？',
'answers': [{'id': '1', 'text': '當梵語發展成俗語之後', 'answer_start': 42},
{'id': '2', 'text': '當梵語發展成俗語之後', 'answer_start': 42}]}],
'id': '1147-6-2',
'question': '天城文在何時成為梵語的標準書寫系統？',
'answers': [{'id': '1', 'text': '19世紀晚期', 'answer_start': 115},
{'id': '2', 'text': '19世紀晚期', 'answer_start': 115}]}],
'id': '1147-6-3',
'question': '最初梵語以什麼書寫系統被記錄下來？',
'answers': [{'id': '1', 'text': '婆羅米文', 'answer_start': 201},
{'id': '2', 'text': '婆羅米文', 'answer_start': 201}]}]}

In [5]: pred['1147-6-1']
Out[5]: '當梵語發展成俗語之後'

In [6]: pred['1147-6-2']
Out[6]: '19世紀晚期'

In [7]: pred['1147-6-3']
Out[7]: '婆羅米文'
```

## 1.(b) Incorrectly

```
In [16]: dev['data'][381]['paragraphs'][0]
Out[16]:
{'context': '2014年世界盃外圍賽，韓國在首輪分組賽以首名出線次輪分組賽，與伊朗、卡達、烏茲別克以及黎巴嫩爭逐兩個直接出線決賽周資格，最後韓國僅以較佳的得失球差壓倒烏茲別克，以小組次名取得2014年世界盃決賽周參賽資格，也是韓國連續八次晉身世界盃決賽周。可惜南韓在決賽周表現不濟，三戰一和兩負小組末席出局。2018年世界盃外圍賽，韓國再次在首輪分組賽以首名出線次輪分組賽，再與伊朗、卡達、烏茲別克同組，同組還有中國及敘利亞。最後韓國以兩分壓倒敘利亞及烏茲別克，再以小組次名取得2018年世界盃決賽周參賽資格，也是韓國連續九次晉身世界盃決賽周。韓國的世界盃成績雖然是亞洲最佳，但在亞洲盃足球賽成績就遠不如世界盃。韓國除了在首兩屆亞洲杯奪冠外，但之後一直與亞洲盃銅牌無緣，自1992年至2011年更連續六屆未能打入過亞洲盃決賽。2015年亞洲盃足球賽，韓國以五連勝一球不失的姿態，廿七年來首次打入亞洲盃決賽，對手是東道主澳洲。雖然韓國在分組初賽曾以1-0擊敗澳洲，但這場決賽韓國卻先失一球，最後在下半場補時階段扳平，令比賽進入加時階段，可惜澳洲最後在加時階段攻入致勝一球，最後韓國以1-2敗陣，只得亞軍。',
'id': '6518-2',
'qas': [{'id': '6518-2-1',
'question': '韓國在哪一年的時候第八次晉身世界盃決賽周?',
'answers': [{'id': '1', 'text': '2014年', 'answer_start': 88},
{'id': '2', 'text': '2014', 'answer_start': 88}]},
{'id': '6518-2-2',
'question': '2018年世界盃外圍賽中，韓國打敗兩個國家的代表隊，取得2018年世界盃決賽周參賽資格，其中一個國家是敘利亞，另一個國家是?',
'answers': [{'id': '1', 'text': '烏茲別克', 'answer_start': 219},
{'id': '2', 'text': '烏茲別克', 'answer_start': 219}]},
{'id': '6518-2-3',
'question': '哪一個國家負責舉辦2015年亞洲盃足球賽?',
'answers': [{'id': '1', 'text': '澳洲', 'answer_start': 402},
{'id': '2', 'text': '澳洲', 'answer_start': 402}]}]}

In [17]: pred['6518-2-2']
Out[17]: '中國'
```

## 2 Which hyperparameter(s) should be modified in order to reach better performance? (e.g. learning\_rate, batch\_size, warmup\_steps, layer\_norm\_eps, attention\_probs\_dropout\_prob)

### 2.(a) Hyperparameters

warmup\_steps should be modified to reach better performance in Kaggle. However, the modification did not perform better in developing-set.

Kaggle	Eval	learning_rate	warmup_steps	atten_dropout
0.75177	<b>81.89557</b>	1.00E-05	0	0.1
0.75979	81.75369	1.00E-05	10	0.1
0.76265	81.18615	1.00E-05	50	0.1
<b>0.76380</b>	81.44154	1.00E-05	200	0.1

In addition, I try to modified learning\_rate and attention\_probs\_dropout\_prob, but they did not perform better.

Kaggle	Eval	learning_rate	warmup_steps	atten_dropout
0.71477	76.41884	1.00E-04	10	0.1
<b>0.75979</b>	<b>81.75369</b>	1.00E-05	10	0.1
0.68762	74.20545	1.00E-06	10	0.1
-	81.21453	1.00E-05	200	0
<b>0.76380</b>	<b>81.44154</b>	1.00E-05	200	0.1
0.75807	81.15778	1.00E-05	200	0.2

## 2.(b) Why it works?

我認為調整 warm up steps 會有用的原因可能是因為 transformers 提供的 bert-base-chinese 使用的中文資料不夠多，所以在 warm up steps 太小的時候比較容易 overfit 到 training data。所以調大了 warm up steps 可以讓 Kaggle 分數上升。但是其實 developing-set 的分數並沒有跟著上升，這部分可能是因為 developing-set 跟 training-set 比較像，testing-set 跟 training-set 比較不像，或是也有可能只是因為 randomness 的影響。