

1. **(2%) After your model predicts the probability of answer span start/end position, what rules did you apply to determine the final start/end position?**

- 若有一組候選機率高於原先被選到最大的機率，則要先確定 start_index 是否小於等於 end_index，且此候選答案字串的長度是否小於 50 字，兩條件都滿足才使用該機率進行更新。

- 若選出來的最終答案中有 [UNK]，則盡量將 [UNK] 試著對應回原先文本中該位置的原始 Token，並將答案更新，確保選出來的答案中不會有 [UNK]。

- 若選出來的最終答案中有 "「" 出現在結尾，且答案中沒有出現 "「"，則將結尾的該引號移除。

2. **(2%) Try another type of pretrained model which can be found in huggingface's Model Hub, and describe**

- the pretrained model you used
- performance of the pretrained model you used
- the difference between BERT and the pretrained model you used

- 使用的 Pretrained Model : luhua/chinese_pretrain_mrc_macbert_large

- 表現高過 bert-base-chinese 不少，在進行最終參數最佳化前，更換至該模型後 Accuracy 表現約提升 5%，最終在 Public Leaderboard 成績是 84.59%

- 差異如下

1. MacBERT 比 BERT 更進一步使用 Whole Word Masking + N-gram Masking 來遮蔽單詞

2. 由於 [MASK] Token 不會在 Fine-tune 階段出現，因此為了避免 Pre-train 與 Fine-tune 的差異，在 Pre-train 時 MacBERT 是使用隨機選擇的同義詞對目標 Token 進行替換遮蔽

3. MacBERT 使用「預測句子順序 (Sentence Order Prediction)」任務取代 BERT 原本的「預測是否為下一句 (Next Sentence Prediction)」任務