

# Machine Learning

## HW7

R10921104 王怡堯

May 5, 2022

### 1 After your model predicts the probability of answer span start/end position, what rules did you apply to determine the final start/end position?

In my first approach, I find the **maximal** start position before **maximum** end position, and the **maximal** end position before **maximum** start position, then determine whether it's greater than max prob or not. Note that the case which maximum start position less than the maximum end position will be include in both case I determined.

In my second approach, I use prefix max to find the maximum probability pair where start index is less than end index. Using a simple Dynamic Programming Algorithm. The performance increased.

### 2 Try another type of pretrained model which can be found in hugging-face's Model Hub, and describe

#### 2.1 the pretrained model you used

macbert and Roberta-wwm (ensemble)

#### 2.2 performance of the model you used

0.84913

#### 2.3 the difference between BERT and the pretrained model you used

Roberta : Dynamic Masking

wwm : Whole Word Masking

Macbert : MLM as correction