

ML2022-2023 Spring HW07 Report

Report Questions

Question 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? (the rules you applied must be different from the sample code)

Answer:

In the sample code, when selecting a paragraph, the answer was always in the middle of the paragraph. This would cause the model to learn strange things. For example, the answers output during testing were more towards the center. This is obviously wrong, because the test We don't know where the answer will end up, so when selecting a paragraph, we should randomly select the starting position from the leftmost to the rightmost.

Question 2

Try another type of pretrained model which can be found in huggingface's Model Hub (e.g. BERT -> BERT-wwm-ext, or BERT -> RoBERTa), and describe

- the pretrained model you used
- performance of the pretrained model you used
- the difference between BERT and the pretrained model you used (architecture, pretraining loss, etc.)

Answer:

Pretrained Model: [luhua/chinese_pretrain_mrc_macbert_large](#), which performs better than the origin BERT model. MacBERT performs better than the traditional BERT model on Chinese tasks by introducing improved masking strategies (such as whole-word masking and synonym replacement masking) and optimizing pre-training objectives. Both remain consistent in their underlying architecture, but are optimized in detail and pre-training methods to suit the needs of specific languages.