bert

geo

2022-04-25

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)

Text

Description automatically generated

1. 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 model you used
* the difference between BERT and the pretrained model you used (architecture, pretraining loss, etc.)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| model | architecture | training loss | traning acc | validation acc |
| bert-base-chinese | 11 層 bertLayer , dropout 0.1 , Liner | 0.595 | 0.751 | 0.63 |
| uer/roberta-base-chinese-extractive-qa | 11 層 bertLayer , dropout 0.1 , Liner(與上式一模一樣) | 0.491 | 0.775 | 0.591 |