temp

gaoyinghan1997

March 2021

1 Training tinybert

1.1 General Distillation

We don't have to do this since general distilled model are given

1.2 data augment

To run data_augmentation.py, we need to prepare two things
1.glove embedding of corpus.
I obtained glove embedding of CoLA by using stanford glove(https://nlp.stanford.edu/projects/glove/)
which worked fine.
2.pre-trained bert(BERT-Base, Uncased) you can find the model here https://github.com/google-research/bert

1.3 Task-specific Distillation

We need a fine-tuned Bert model as the teacher model.(Which I am working for now)
1.For Fine-tuning with BLeU(classification tasks) we need to rewrite 'processor'
Class in 'run_classifier.py', which should be easy. Where I am stuck now is to
run 'run_classifier.py' I had syntax errors when I was using shell to run it. I
tried configure it manually but still got error:
class AdamWeightDecayOptimizer(tf.train.Optimizer):
AttributeError: module 'tensorflow._api.v2.train' has no attribute 'Optimizer'
2.For Fine-tuning with SQuAD, we need to run 'run_squad.py'
Instruction for fine-tuning can be found in https://github.com/google-research/bert

1.4 evaluation

should be straight forward.