

A6 BERT Pretraining

Name: Ziheng Wang

SID: 12310401

Result

```
(DL312) ~/Documents/Sophomore_Second/NLP/Ass6/code/ [main*] python train.py
5503 batch groups loaded
Epoch 0/1, Batch 5040/5503, Loss 5.88, Remaining 130.72 sec for current epoch^R[A
Epoch 0/1, Batch 5500/5503, Loss 7.53, Remaining 0.85 sec for current epochhh
```

Evaluation

```
(DL312) ~/Documents/Sophomore_Second/NLP/Ass6/code/ [main*] python eval.py bert_model.pt
NSP Accuracy: 0.5000
Mask Accuracy: 0.0937
Total samples: 768
Total mask tokens: 3478
```