1.

python vulrepair\_main.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--tokenizer\_name=Salesforce/codet5-base \

--model\_name\_or\_path=Salesforce/codet5-base \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 4  


2.

python vulrepair\_main.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--tokenizer\_name=Salesforce/codet5-base \

--model\_name\_or\_path=Salesforce/codet5-base \

--model\_bin\_path=./saved\_models/checkpoint-best-loss/model.bin \ 换本地的训练得到的模型权重

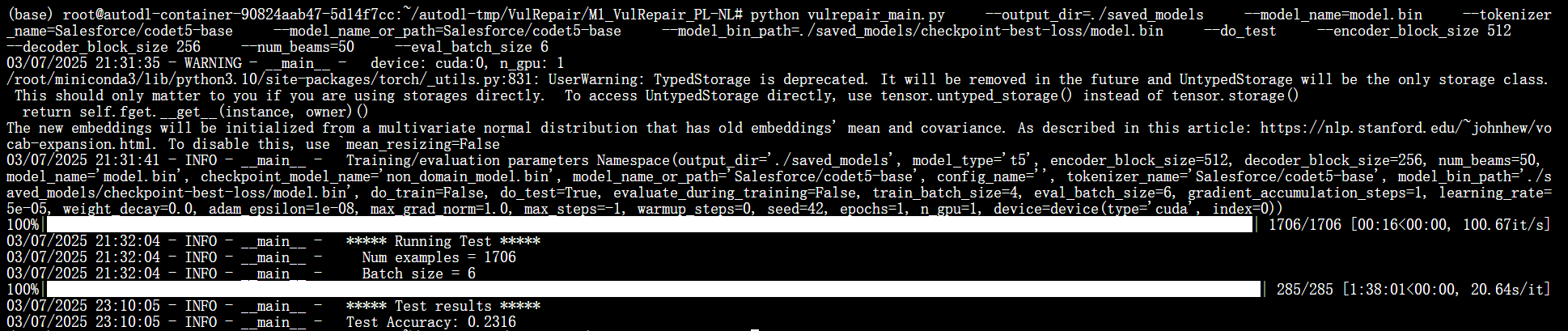
--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 4



？？？？？

3.

python vulrepair\_main.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=MickyMike/VulRepair \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

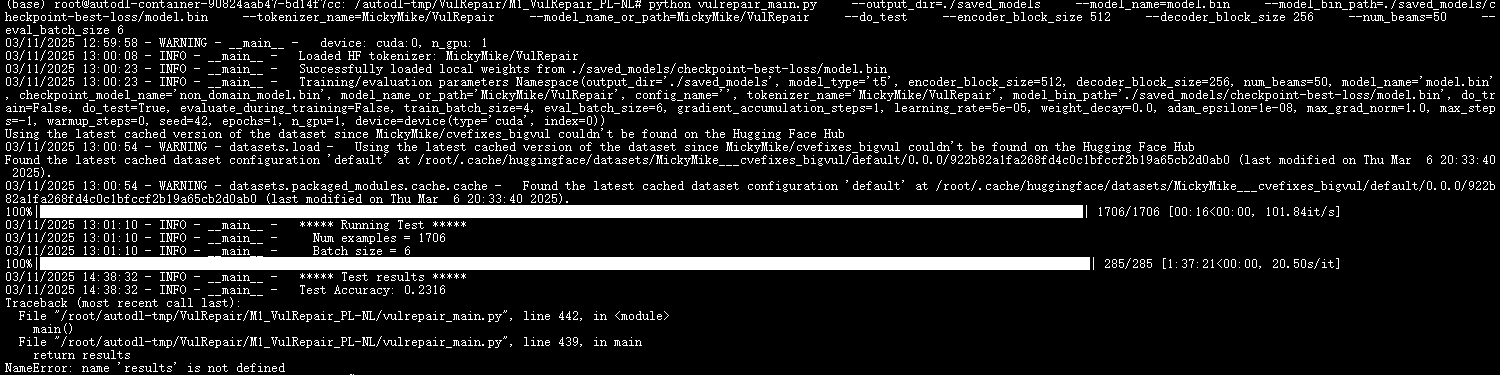
--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 6

使用mike的分词器和配置，本地模型权重，测出准确率仍然是23.16



4.

用自己重新训练的模型权重测试结果：  
python vulrepair\_main.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./Inference\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=MickyMike/VulRepair \

--model\_name\_or\_path=MickyMike/VulRepair \

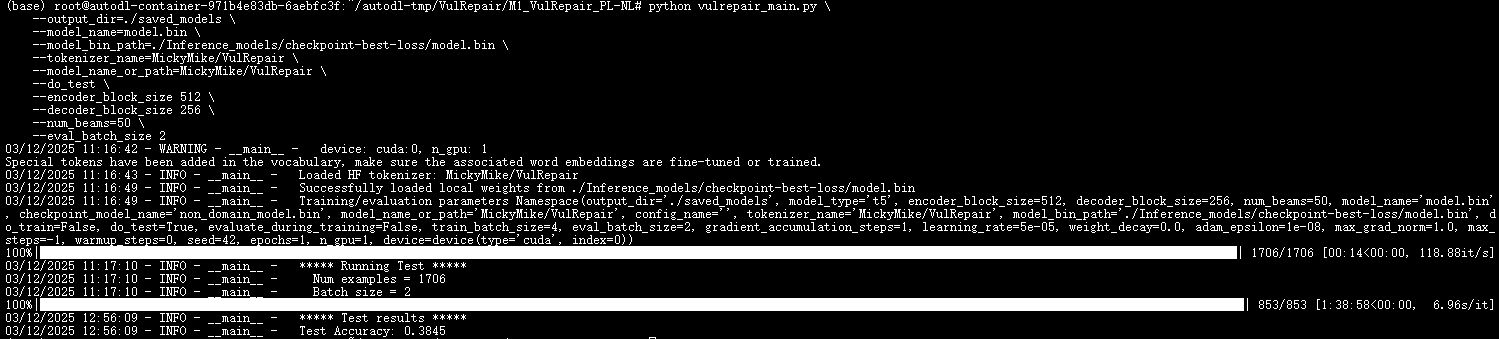
--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 2



5.

python vulrepair\_main.py \

--output\_dir=./MickyVulRepair\_models \

--model\_name=model.bin \

--model\_weights\_dir=/root/autodl-tmp/VulRepair/M1\_VulRepair\_PL-NL/codet5-base\_models/saved\_model\_weight/2025-03-17-19h \

--model\_path=/root/autodl-tmp/VulRepair/models/MickyVulRepair \

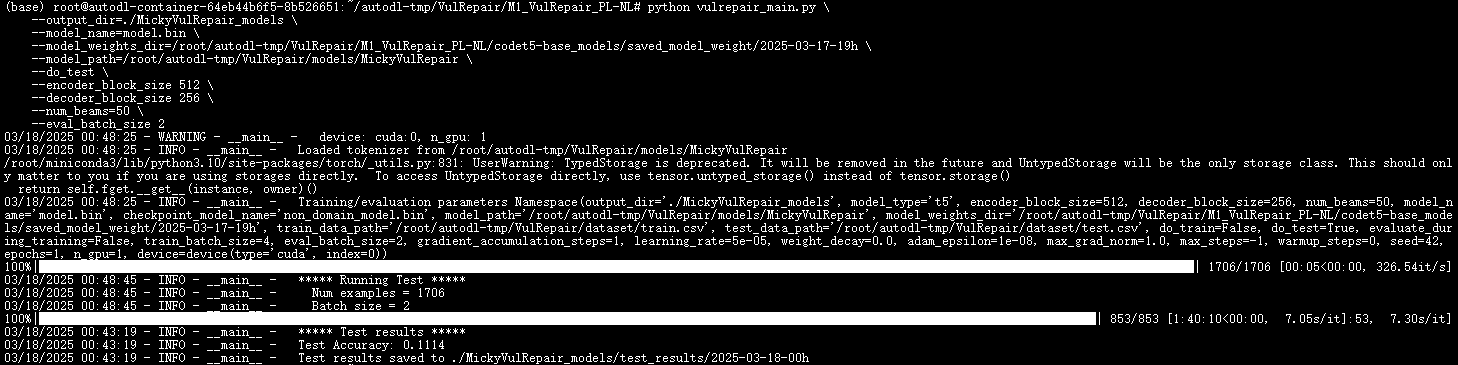
--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 2



6.

python vulrepair\_main.py \

--output\_dir=./MickyVulRepair\_models \

--model\_name=model.bin \

--model\_weights\_dir=/root/autodl-tmp/VulRepair/M1\_VulRepair\_PL-NL/codet5-base\_models/saved\_model\_weight/2025-03-17-19h \

--model\_path=/root/autodl-tmp/VulRepair/models/MickyVulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 2



7.

再跑一下

python vulrepair\_main.py \

--output\_dir=./local\_models \

--model\_name=model.bin \

--model\_weights\_dir=/root/autodl-tmp/VulRepair/M1\_VulRepair\_PL-NL/local\_models \

--model\_path=/root/autodl-tmp/VulRepair/models/MickyVulRepair \

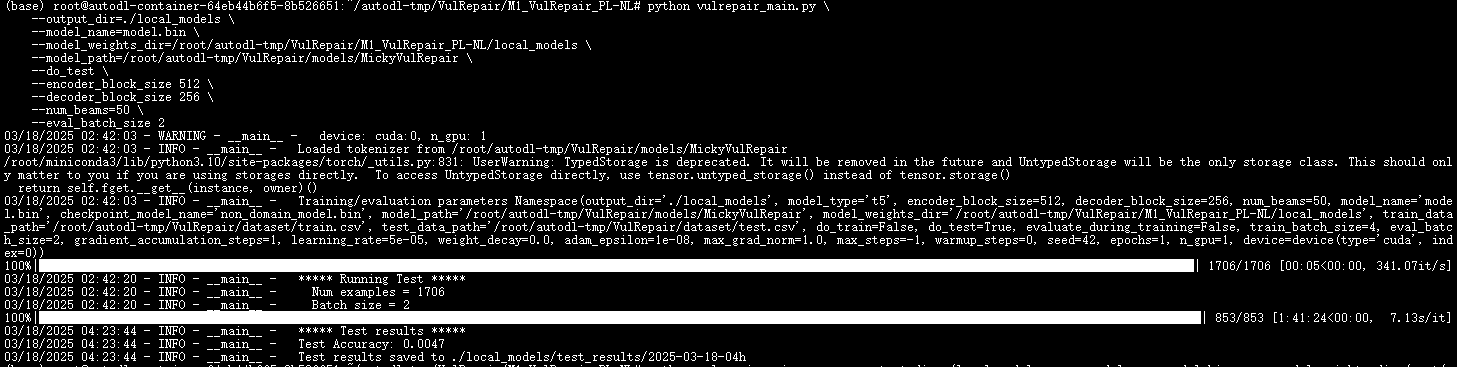
--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 2



8.用micky的model.bin测试

python vulrepair\_main.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./local\_models/pytorch\_model.bin \

--tokenizer\_name=MickyMike/VulRepair \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 2

03/19/2025 14:31:50 - INFO - \_\_main\_\_ - \*\*\*\*\* Running Test \*\*\*\*\*

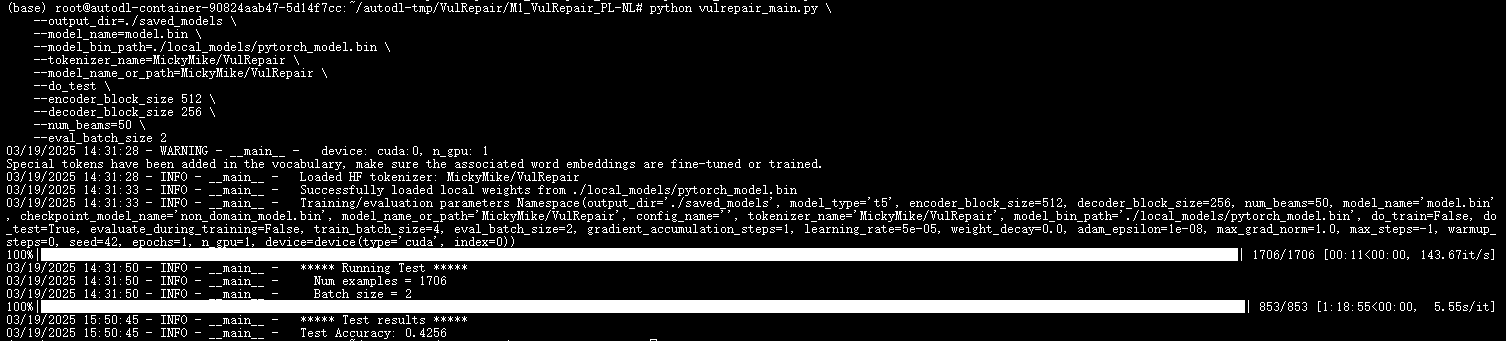
03/19/2025 14:31:50 - INFO - \_\_main\_\_ - Num examples = 1706

03/19/2025 14:31:50 - INFO - \_\_main\_\_ - Batch size = 2

100%|████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████| 853/853 [1:18:55<00:00, 5.55s/it]

03/19/2025 15:50:45 - INFO - \_\_main\_\_ - \*\*\*\*\* Test results \*\*\*\*\*

03/19/2025 15:50:45 - INFO - \_\_main\_\_ - Test Accuracy: 0.4256



9.终于训练出比micky准确率还高的模型权重

python vulrepair\_main.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./Inference\_3.19-16h\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=MickyMike/VulRepair \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

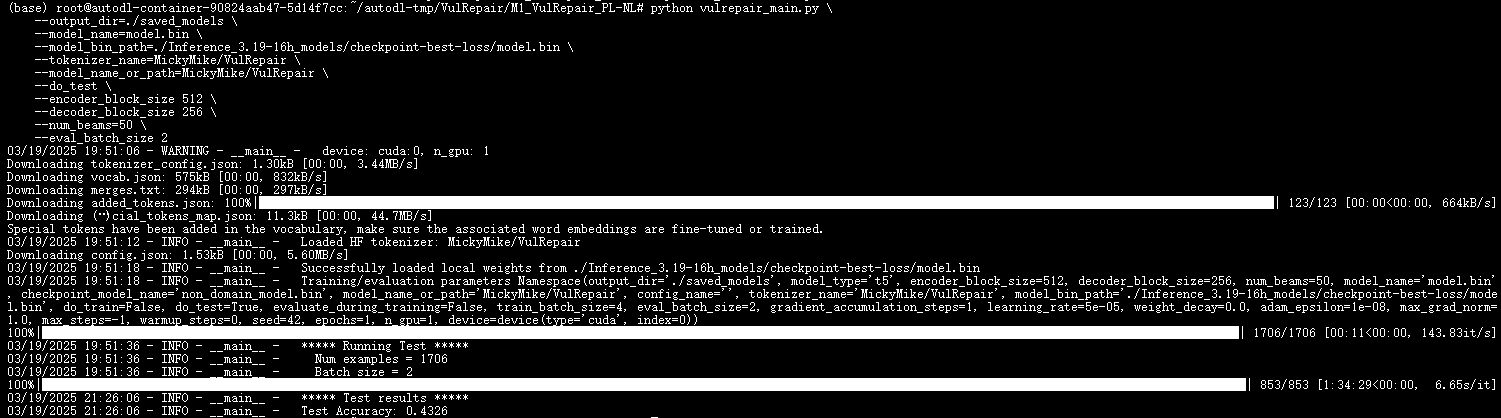
--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 4

这个不错



9.服务器测试43.26model.bin测试结果是多少

python vulrepair\_main.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./model\_bin/pytorch\_model.bin \

--tokenizer\_name=Salesforce/codet5-base \

--model\_name\_or\_path=Salesforce/codet5-base \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 8

python vulrepair\_main\_copy.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./Inference\_3.24-18h\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=MickyMike/VulRepair \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 4  
以上命令测出来24左右准确率，今天训练的，训练写的40轮，但是在29轮左右自动停止，是loss 三轮没有下降就停止的代码。

10. 用micky数据集43.26准确率的model.bin再训练数据集semeru/code-code-BugFixingMed：

python semeru.py \

--model\_name=model.bin \

--output\_dir=./Inference\_models \

--tokenizer\_name=Salesforce/codet5-base \

--model\_name\_or\_path=Salesforce/codet5-base \

--model\_bin\_path=./model\_bin/0\_4326model.bin \

--do\_train \

--epochs 15 \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--train\_batch\_size 16 \

--eval\_batch\_size 16 \

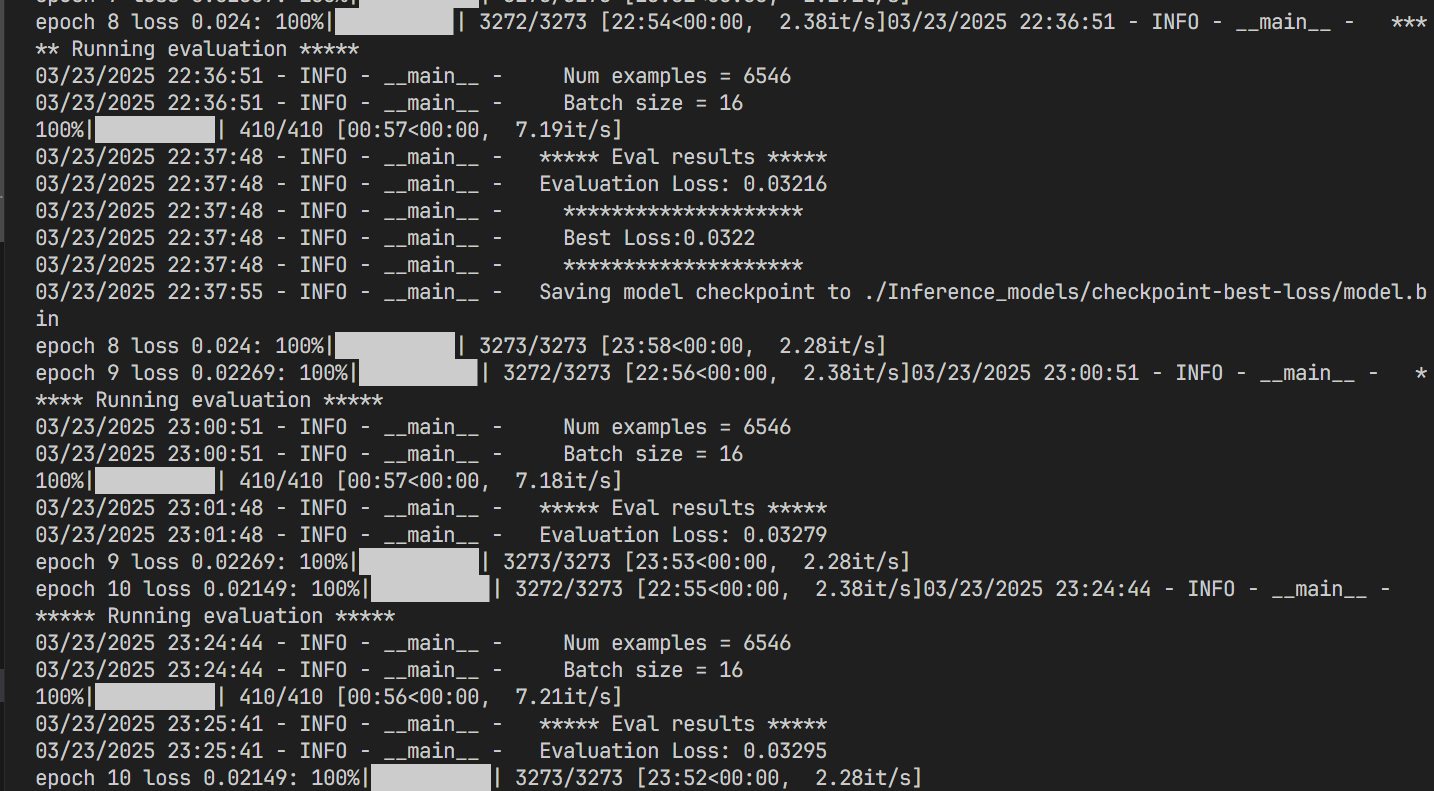
--learning\_rate 2e-5 \

--max\_grad\_norm 1.0 \

--evaluate\_during\_training \

--seed 123456 \

2>&1 | tee train.log



再进行测试：

python semeru.py \

--model\_name=model.bin \

--output\_dir=./Inference\_models \

--tokenizer\_name=Salesforce/codet5-base \

--model\_name\_or\_path=Salesforce/codet5-base \

--model\_bin\_path=./Inference\_models/checkpoint-best-loss/model.bin \

--do\_test \

--encoder\_block\_size 512 \

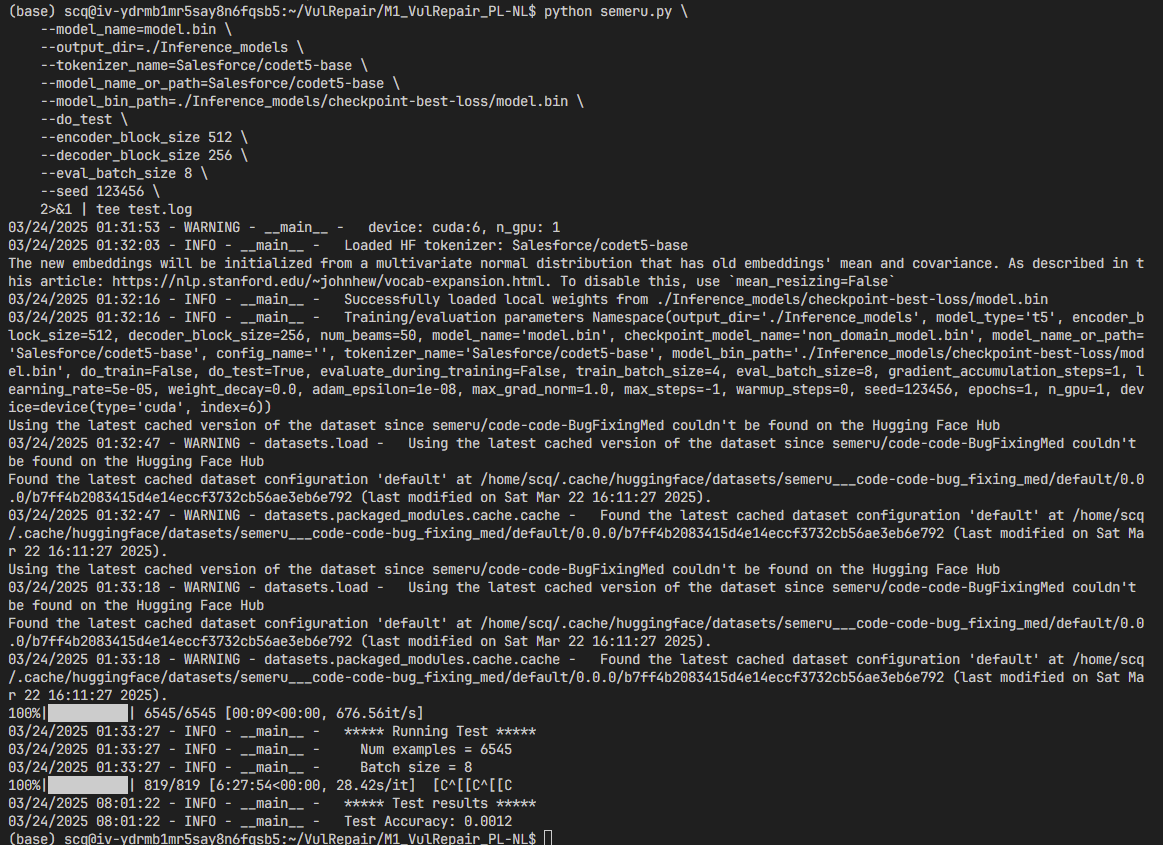
--decoder\_block\_size 256 \

--eval\_batch\_size 8 \

--seed 123456 \

2>&1 | tee test.log

WTF?



11.很好，使用vulrepair\_main\_3-25.py

python vulrepair\_main.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./03-25-23h29m\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=./saved\_models/tokenizer \

--model\_name\_or\_path=MickyMike/VulRepair \

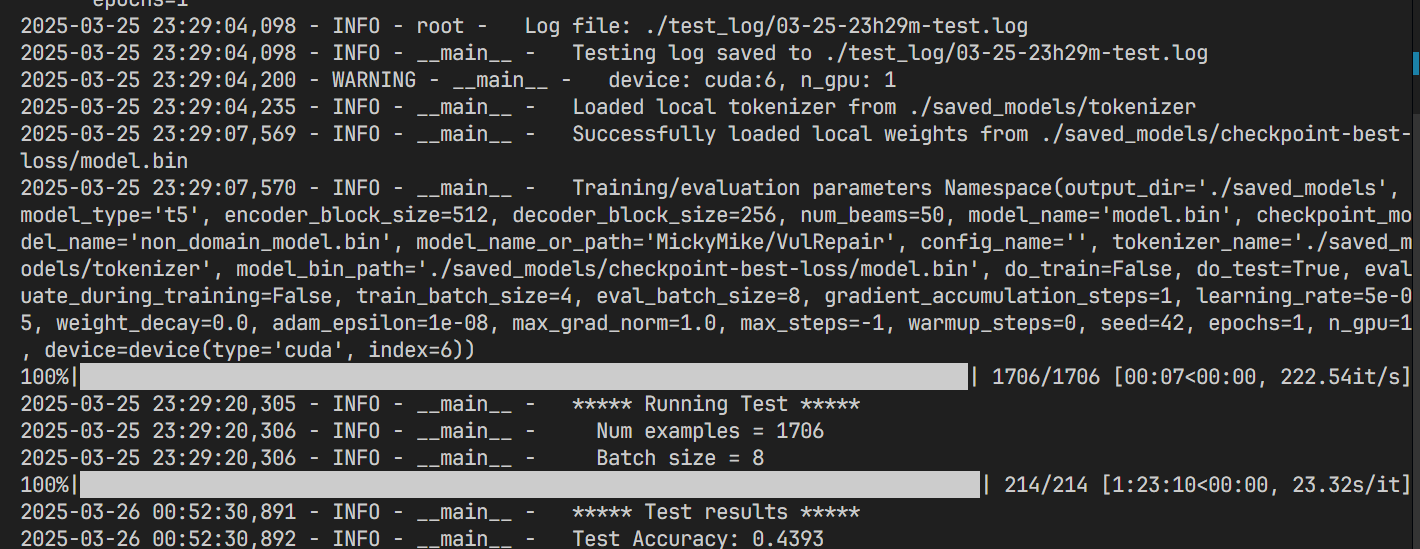
--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 8



12. 使用最新的vulrepair\_main\_3-28.py来训练，增加了各种参数，结果更差了

python vulrepair\_main\_3-28.py \

--model\_name=model.bin \

--output\_dir=./3-29-20h\_saved\_models \

--tokenizer\_name=Salesforce/codet5-base \

--model\_name\_or\_path=Salesforce/codet5-base \

--do\_train \

--epochs 40 \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--train\_batch\_size 16 \

--eval\_batch\_size 16 \

--learning\_rate 2e-5 \

--max\_grad\_norm 1.0 \

--evaluate\_during\_training \

--seed 123456 \

--fp16 \

--label\_smoothing 0.1 \

--top\_p 0.95 \

--top\_k 50 \

--temperature 0.7 2>&1 | tee train.log

训练结果：  
python vulrepair\_main\_3-28.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./3-29-20h\_saved\_models/checkpoint-best/model.bin \

--tokenizer\_name=./3-29-20h\_saved\_models/tokenizer \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

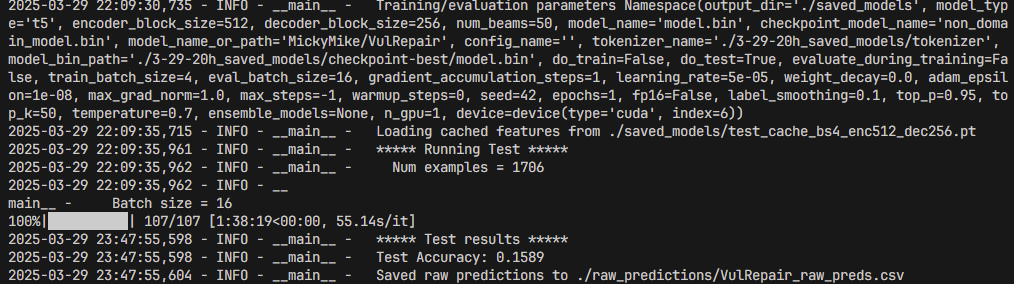
--num\_beams 50 \

--eval\_batch\_size 16 \

--top\_p 0.95 \

--top\_k 50 \

--temperature 0.7 2>&1 | tee test.log



13.更改方案，修改vulrepair\_main\_3-25.py使其评估方式变为部分匹配或者语义相同等方式来提高准确率

python vulrepair\_main\_3-28.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./03-25-23h29m\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=./03-25-23h29m\_saved\_models/tokenizer \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

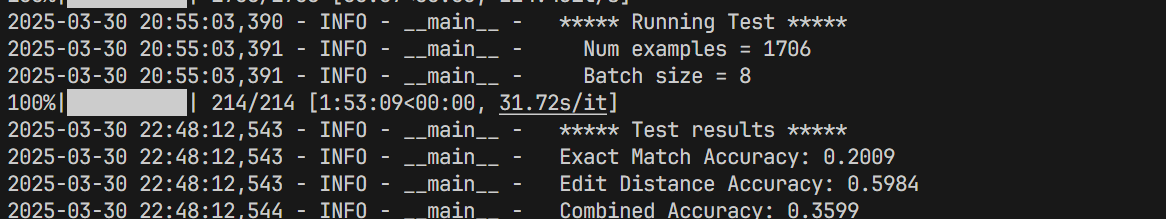
--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 8 \

2>&1 | tee test.log  
奇怪用的11的model.bin但是准确率下降好多



14.再试下3-25，跟13一样，和11又不一样了。就离谱

python vulrepair\_main\_3-25.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./03-25-23h29m\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=./03-25-23h29m\_saved\_models/tokenizer \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

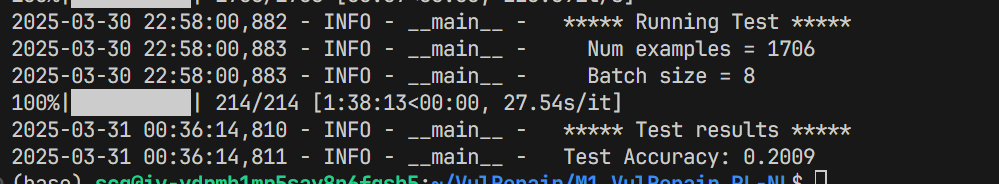
--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 8 \

2>&1 | tee test.log



15.tmux里面换个分词器试试

python vulrepair\_main\_3-25.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./03-25-23h29m\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=MickyMike/VulRepair \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 8 \

2>&1 | tee test.log

跟上面一样

16.重新训练 结果仍然跟上面一样

（1）

训练：

python vulrepair\_main\_3-25.py --model\_name=model.bin --output\_dir=./3-31-10h\_main\_3-25\_saved\_models --tokenizer\_name=Salesforce/codet5-base --model\_name\_or\_path=Salesforce/codet5-base --do\_train --epochs 40 --encoder\_block\_size 512 --decoder\_block\_size 256 --train\_batch\_size 16 --eval\_batch\_size 16 --learning\_rate 2e-5 --max\_grad\_norm 1.0 --evaluate\_during\_training --seed 123456 2>&1 | tee train.log

测试：

python vulrepair\_main\_3-25.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./3-31-10h\_main\_3-25\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=./3-31-10h\_main\_3-25\_saved\_models/tokenizer \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 8

（2）

训练：

python vulrepair\_main\_3-30.py --model\_name=model.bin --output\_dir=./3-31-10h\_main\_3-30\_saved\_models --tokenizer\_name=Salesforce/codet5-base --model\_name\_or\_path=Salesforce/codet5-base --do\_train --epochs 40 --encoder\_block\_size 512 --decoder\_block\_size 256 --train\_batch\_size 16 --eval\_batch\_size 16 --learning\_rate 2e-5 --max\_grad\_norm 1.0 --evaluate\_during\_training --seed 123456 2>&1 | tee train.log

测试：

python vulrepair\_main\_3-30.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=./3-31-10h\_main\_3-30\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=./3-31-10h\_main\_3-30\_saved\_models/tokenizer \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 8

17.

训练：

python vulrepair\_main\_3-30.py \

--model\_name=model.bin \

--output\_dir=/data/share/data/scq/vulrepair/m1/4-2-19h\_saved\_models \

--tokenizer\_name=Salesforce/codet5-base \

--model\_name\_or\_path=Salesforce/codet5-base \

--do\_train \

--epochs 40 \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--train\_batch\_size 8 \

--eval\_batch\_size 8 \

--learning\_rate 2e-5 \

--max\_grad\_norm 1.0 \

--evaluate\_during\_training \

--seed 123456 \

2>&1 | tee train.log

测试：

（1）

python vulrepair\_main\_3-30.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=/data/share/data/scq/vulrepair/m1/4-2-19h\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=/data/share/data/scq/vulrepair/m1/4-2-19h\_saved\_models/tokenizer \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 16

（2）

python vulrepair\_main\_3-30.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=/data/share/data/scq/vulrepair/m1/4-2-19h\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=/data/share/data/scq/vulrepair/m1/4-2-19h\_saved\_models/tokenizer \

--model\_name\_or\_path=Salesforce/codet5-base \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 16

18.无敌

python vulrepair\_main\_3-30.py \

--model\_name=model.bin \

--output\_dir=../saves/m1/4-3-20h\_saved\_models \

--tokenizer\_name=Salesforce/codet5-base \

--model\_name\_or\_path=Salesforce/codet5-base \

--do\_train \

--epochs 40 \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--train\_batch\_size 8 \

--eval\_batch\_size 8 \

--learning\_rate 2e-5 \

--max\_grad\_norm 1.0 \

--evaluate\_during\_training \

--seed 123456 \

2>&1 | tee train.log

测试：

python vulrepair\_main\_3-30.py \

--output\_dir=./saved\_models \

--model\_name=model.bin \

--model\_bin\_path=../saves/m1/4-3-20h\_saved\_models/checkpoint-best-loss/model.bin \

--tokenizer\_name=../saves/m1/4-3-20h\_saved\_models/tokenizer \

--model\_name\_or\_path=MickyMike/VulRepair \

--do\_test \

--encoder\_block\_size 512 \

--decoder\_block\_size 256 \

--num\_beams=50 \

--eval\_batch\_size 8

