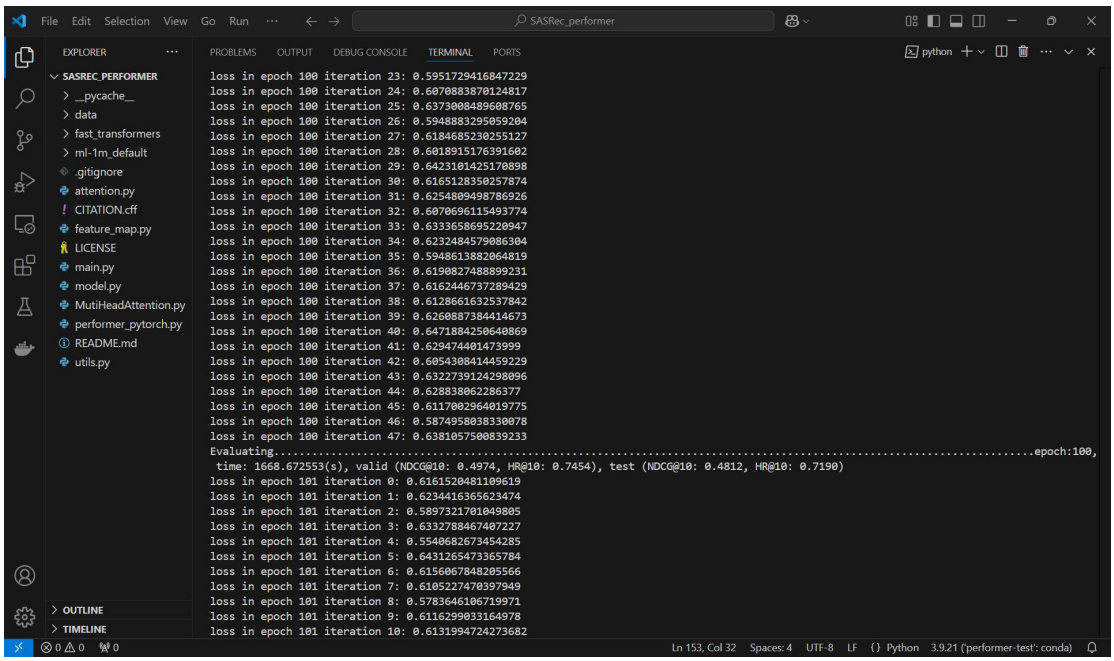
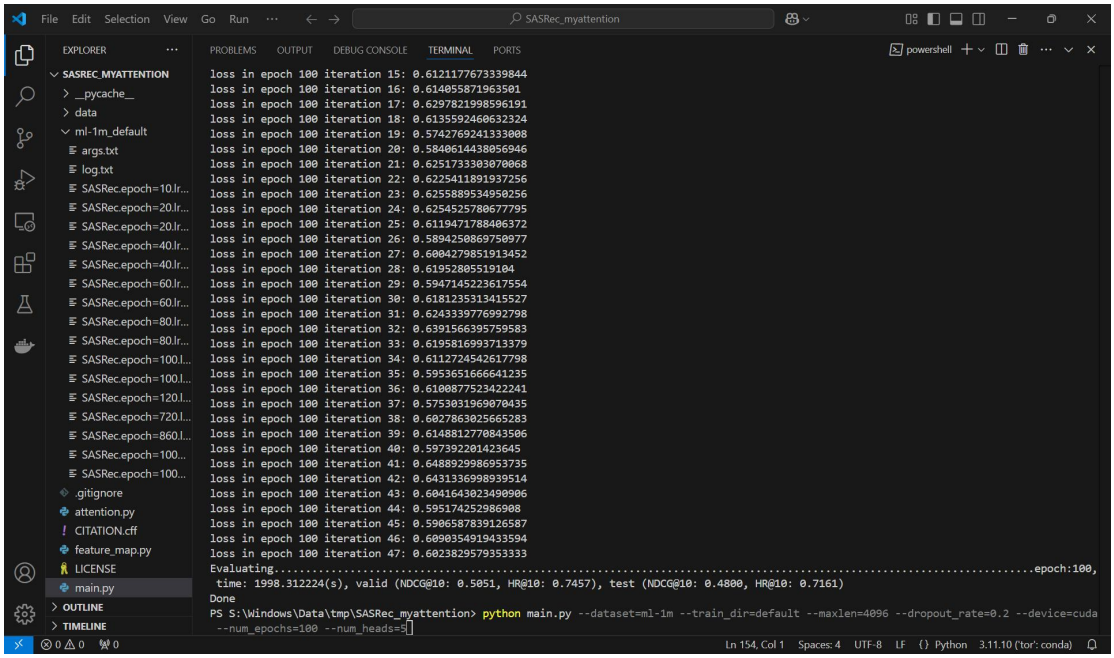
Performer

python main.py --dataset=ml-1m --train_dir=default --maxlen=4096 --dropout_rate=0.2 --device=cuda --num_epochs=100 --num_heads=5



Myattention

python main.py --dataset=ml-1m --train_dir=default --maxlen=4096 --dropout_rate=0.2 --device=cuda --num_epochs=100 --num_heads=5



时间 1998->1668

优化了 16.5%，性能几乎不变。