

1 目录结构

/src 源代码所在位置

/src/main.py 主体的训练，测试代码

/src/main_nlp.py 语言建模的训练，测试代码

/src/constants.py 常量

/src/plot.py 结果可视化相关代码

/src/algorithm 我们自己实现的RAdam和lookahead

/src/data 几种不同数据集的数据读取和预处理位置

/src/models 几种不同问题对应模型的读取位置

/results 实验结果所在位置

/results/text 实验结果的txt文档

/results/runs 实验结果的tensorboard可视化文件

/results/picture 实验图片结果

/data 实验数据所在位置

/data/cifar-10-batches-py cifar10数据

/data/cifar-100-python cifar100数据

/data/cnn_data 作业2的部分ImageNet图片数据

/data/ptb 作业3的语言建模数据

IMDB数据会自动下载

2 实验环境

- ubuntu 18.04系统
- python 3.6.9
- pytorch 1.4.0
- numpy 1.18.4
- torchvision 0.5.0
- tensorboard 1.14.0
- keras 2.3.1

3 实验复现方法说明

3.1 基本性能比较

实验复现方法

(注：以下实验均运行在src目录下)

cifar10:

Ranger(RAdam + lookahead):

```
python main.py
```

RAdam:

```
python main.py --lookahead 0
```

cifar100:

Ranger(RAdam + lookahead):

```
python main.py --dataset cifar100
```

RAdam:

```
python main.py --dataset cifar100 --lookahead 0
```

可视化复现方法

(注：以下可视化均运行在results目录下)

cifar10的train loss对比

Ranger(RAdam + lookahead):

```
tensorboard --logdir runs/000
```

cifar10的test accuracy对比

```
tensorboard --logdir runs/001
```

cifar100的train loss对比

```
tensorboard --logdir runs/100
```

cifar100的test accuracy对比

```
tensorboard --logdir runs/101
```

3.2 对超参数鲁棒

实验复现方法

(注：以下实验均运行在src目录下)

cifar10:

Ranger (标准超参数 $\alpha = 0.8$, $k = 5$)

```
python main.py
```

Ranger ($\alpha = 0.5$)

```
python main.py --lookahead_lr 0.5
```

Ranger (k = 10)

```
python main.py --lookahead_steps 10
```

RAdam:

```
python main.py --lookahead 0
```

cifar100:

Ranger (标准超参数 $\alpha = 0.8$, $k = 5$)

```
python main.py --dataset cifar100
```

Ranger ($\alpha = 0.5$)

```
python main.py --lookahead_lr 0.5 --dataset cifar100
```

Ranger (k = 10)

```
python main.py --lookahead_steps 10 --dataset cifar100
```

RAdam:

```
python main.py --lookahead 0 --dataset cifar100
```

可视化复现方法

(注：以下可视化均运行在results目录下)

cifar10的train loss对比

```
tensorboard --logdir runs/030
```

cifar10的test accuracy对比

```
tensorboard --logdir runs/031
```

cifar100的train loss对比

```
tensorboard --logdir runs/130
```

cifar100的test accuracy对比

```
tensorboard --logdir runs/131
```

3.3 对学习率鲁棒

3.3.1 学习率变化

实验复现方法

(注：以下实验均运行在src目录下)

cifar10:

Ranger(RAdam + lookahead):

```
python main.py --learning_rate 0.01
```

RAdam:

```
python main.py --learning_rate 0.01 --lookahead 0
```

cifar100:

Ranger(RAdam + lookahead):

```
python main.py --learning_rate 0.01 --dataset cifar100
```

RAdam:

```
python main.py --learning_rate 0.01 --dataset cifar100 --lookahead 0
```

可视化复现方法

(注：以下可视化均运行在results目录下)

cifar10的train loss对比

```
tensorboard --logdir runs/010
```

cifar10的test accuracy对比

```
tensorboard --logdir runs/011
```

cifar100的train loss对比

```
tensorboard --logdir runs/110
```

cifar100的test accuracy对比

```
tensorboard --logdir runs/111
```

3.3.2 学习率衰减方式变化

实验复现方法

(注：以下实验均运行在src目录下)

cifar10:

Ranger(RAdam + lookahead):

```
python main.py --lr_decay fast
```

RAdam:

```
python main.py --lr_decay fast --lookahead 0
```

cifar100:

Ranger(RAdam + lookahead):

```
python main.py --lr_decay fast --dataset cifar100
```

RAdam:

```
python main.py --lr_decay fast --dataset cifar100 --lookahead 0
```

可视化复现方法

(注：以下可视化均运行在results目录下)

cifar10的train loss对比

```
tensorboard --logdir runs/020
```

cifar10的test accuracy对比

```
tensorboard --logdir runs/021
```

cifar100的train loss对比

```
tensorboard --logdir runs/120
```

cifar100的test accuracy对比

```
tensorboard --logdir runs/121
```

3.4 Lookahead + other optimizer

实验复现方法

(注：以下实验均运行在src目录下)

cifar10:

Ranger(RAdam + lookahead):

```
python main.py
```

RAdam:

```
python main.py --lookahead 0
```

Adam + lookahead:

```
python main.py --algorithm Adam
```

Adam:

```
python main.py --algorithm Adam --lookahead 0
```

SGD + lookahead:

```
python main.py --algorithm SGD
```

SGD:

```
python main.py --algorithm SGD --lookahead 0
```

cifar100:

Ranger(RAdam + lookahead):

```
python main.py --dataset cifar100
```

RAdam:

```
python main.py --lookahead 0 --dataset cifar100
```

Adam + lookahead:

```
python main.py --algorithm Adam --dataset cifar100
```

Adam:

```
python main.py --algorithm Adam --lookahead 0 --dataset cifar100
```

SGD + lookahead:

```
python main.py --algorithm SGD --dataset cifar100
```

SGD:

```
python main.py --algorithm SGD --lookahead 0 --dataset cifar100
```

可视化复现方法

(注：以下可视化均运行在results目录下)

cifar10的train loss对比

```
tensorboard --logdir runs/040
```

cifar10的test accuracy对比

```
tensorboard --logdir runs/041
```

cifar100的train loss对比

```
tensorboard --logdir runs/140
```

cifar100的test accuracy对比

```
tensorboard --logdir runs/141
```

3.5 性能提升是否来源于batch size的提升?

实验复现方法

(注：以下实验均运行在src目录下)

cifar10:

步长5 + batch128

```
python main.py
```

步长2 + batch320

```
python main.py --lookahead_steps 2 --batch_size 320
```

步长10 + batch64

```
python main.py --lookahead_steps 10 --batch_size 64
```

cifar100:

步长5 + batch128

```
python main.py --dataset cifar100
```

步长2 + batch320

```
python main.py --lookahead_steps 2 --batch_size 320 --dataset cifar100
```

步长10 + batch64

```
python main.py --lookahead_steps 10 --batch_size 64 --dataset cifar100
```

可视化复现方法

(注：以下可视化均运行在results目录下)

cifar10的train loss对比

```
tensorboard --logdir runs/050
```

cifar10的test accuracy对比

```
tensorboard --logdir runs/051
```

cifar100的train loss对比

```
tensorboard --logdir runs/150
```

cifar100的test accuracy对比

```
tensorboard --logdir runs/151
```

3.6 其他任务/模型上的测试

3.6.1 hw2 image classification

实验复现方法

(注：以下实验均运行在src目录下)

Ranger

```
python main.py --dataset self_cnn
```

RAdam

```
python main.py --dataset self_cnn --lookahead 0
```

可视化复现方法

(注：以下可视化均运行在results目录下)

train loss对比

```
tensorboard --logdir runs/400
```

test accuracy对比

```
tensorboard --logdir runs/401
```

3.6.2 hw3 language model

实验复现方法

(注：以下实验均运行在src目录下)

Ranger

```
python main_nlp.py
```

RAdam


```
python main_nlp.py --lookahead 0
```

可视化复现方法

(注：以下可视化均运行在results目录下)

train loss对比

```
tensorboard --logdir runs/200
```

test accuracy对比

```
tensorboard --logdir runs/201
```

3.6.3 IMDB classification

实验复现方法

(注：以下实验均运行在src目录下)

Ranger

```
python main.py --dataset imdb
```

RAdam

```
python main.py --lookahead 0 --dataset imdb
```

可视化复现方法

(注：以下可视化均运行在results目录下)

train loss对比

```
tensorboard --logdir runs/300
```

test accuracy对比

```
tensorboard --logdir runs/301
```