

对比实验运行步骤：

总共五个对比实验,分别训练两个数据集：

1.atrank 2.Bisle 3.bpr 4.rnn_att 5.BisIE_mask

amazon和**tianchi**两个文件夹分别为两个数据集的预处理文件

raw_data和**shop_data**为两个数据集的存放文件夹

第一个amazon数据集：

第一步：先对数据集进行预处理

1、amazon数据集

```
1 cd amzon;
2 python 1_convert_pd.py; #该文件可以控制amazon数据量的大小
3 python 2_remap_id.py;
```

2、天池数据集

```
1 cd tianchi
2 python 1_ConversionFile.py;
3 python 2_Convert.py; #此文件夹可以控制天池数据集的数据量的大小
4 python 3_Remap.py;
```

第二步：创建训练和测试数据集，并且开始训练数据集

1、amazon数据集

```
1 cd atrank;
2 python build_dataset_az.py;
3 python train.py;
```

2、天池数据集

```
1 cd atrank;
2 cd tianchi;
3 python build_dataset.py #该文件内可以改变sequence的长度，为了防止数据量过大，难以训练
4 cd ..;
5 python train_new.py;
```

注：每一个对比实验文件夹下，都会有一个**tianchi**文件夹（用于生成天池数据集的训练数据集和测试集），**train.py**（amazon数据集的训练文件），**train_new.py**（天池数据集的训练）。