

# 程序说明与测试报告补充

## 程序文件架构

```
├─ config                                #一些模型的配置文件
│   ├── default.config
│   └─ Music
│       ├── DeepFM.config
│       ├── DeepInterest.config
│       ├── LatentRela.config
│       ├── NCF.config
│       ├── OurConfig.config
│       ├── RelationNet.config
│       └─ xDeepFM.config
├─ config_reader                        #config parser
├─ eval.py                             #模型测试代码
├─ model
│   ├── get_model.py
│   ├── loss.py
│   └─ model                            #baseline模型代码
│       ├── DeepFM.py
│       ├── DeepInterest.py
│       ├── FieldEncoder.py
│       ├── LRMM.py
│       ├── MusicEncoder.py
│       ├── NCF.py
│       ├── UserEncoder.py
│       ├── xDeepFM.py
│       └─ Ours                         #我们的模型代码
│           └─ RelationNetwork.py
├─ work.py                             #train与test所用函数
├─ reader                              #每个模型对数据处理的代码
│   ├── formatter
│   │   ├── basic.py
│   │   ├── DeepFMFormatter.py
│   │   ├── DeepInterestFormatter.py
│   │   ├── MusicFormatter.py
│   │   ├── MusicPairFormatter.py
│   │   ├── NCFFormatter.py
│   │   └─ OurFormatter.py
│   └─ reader.py
├─ README.md
├─ train.py                            #模型训练代码
└─ utils                               #一些函数，包括auc指标等
```

## 运行环境与方法

环境：python3 / pytorch

测试: `python eval.py -c config_path -g gpu_idx -m $trained_model_path`

```
python eval.py -c config/Music/DeepFM.config -g 2 -m /mnt/data5/zzy/hw/MusicRec/model/DeepFM-Final2/model-10.pkl
```

训练: `python train.py -c config_path -g gpu_idx`

```
python train.py -c config/Music/DeepFM.config -g 6
```

## 程序功能介绍

### formatter

formatter用于数据处理, 每个model需要实现对应函数

```
class formatter:
    def __init__(self, config):
        pass
    def check(self, data, config):                #判断该数据可否使用
        pass
    def format(self, data, config, transformer, mode):#给出处理后的数据
        pass
```

### reader

reader.py 进行多线程拉取数据,同时设置对应模型的formatter

### model

get\_model.py 拉取model, 新模型需要在此注册

work.py 定义valid\_net和train\_net, 是训练和测试代码的主体

### config

模型配置文件, 主要包含以下几个:

```
[train] #训练相关参数
batch_size = 256
type_of_loss = cross_entropy_loss
.....
[model] #模型参数
name = DeepFM

hidden_size = 64
deep_layer1= 256
deep_layer2= 128
.....
[reader]#数据拉取参数
```

```
train_reader_num = 5 #训练拉取数据线程数量
valid_reader_num = 5 #测试拉取数据线程数量
.....
[data] #数据路径等
formatter = DeepFM
train_data_path = /home/zzy/hw/musicRec/predata/DataFinal/train
train_file_list =
train_0.json,train_1.json,train_2.json,train_3.json,train_4.json,train_5.json,train_6.json,train
_7.json,train_8.json,train_9.json
user_info_path = /home/zzy/hw/musicRec/predata/user_info_without_moments.json
music_info_path = /home/zzy/hw/musicRec/predata/song_info.json
.....
[valid] #测试输出等
valid_out =True
valid_out_path =/home/zzy/hw/musicRec/MusicRecom/res/deepfm
[output]#模型存放等
model_name = DeepFM-Final3
model_path = /mnt/data5/zzy/hw/MusicRec/model/
.....
```