

# 代码审核规则

1. 选手需要提交训练好的模型，以及能够针对初赛测试集产出预测结果的训练、预测代码。所有文件请打包在 `zip` 压缩包内，对提交文件的组织形式要求如下。
2. 赛题举办方（科大讯飞）将匿名对选手进行审核、查重、复现，如果发现选手代码存在作弊、抄袭或无法复现的情况，则取消排行榜成绩。

## 代码打包规范

### 1、原始数据文件夹 `xfdata/`

选手无需打包科大讯飞提供的竞赛数据文件，我们会把提供的所有原始文件（与官网上的文件和文件名一致）放到 `xfdata/` 文件夹下。需要注意，初始的时候 `xfdata` 文件夹会被清空，然后再放入原始数据文件。

数据结构示例如下：

```
|-- xfdata
# 命名与提供给选手的数据集名称保持一致
|--oppo_breeno_round1_data
    |--train.csv
    |--test.csv
```

### 2、用户数据文件夹 `user_data/`

选手训练好的模型、预测过程中需要生成的中间数据，请放入该文件夹中。文件夹下的子目录、文件名，选手可自行决定。

数据结构示例如下：

```
|-- user_data
    |-- model_data
        |-- model.dat
    |-- tmp_data
        |-- tmp.dat
    |-- 计算的中间文件
    |-- 临时存储目录
```

### 3、预测结果输出文件夹 `prediction_result/`

选手提交的代码，需要在此文件夹中产出针对测试集的预测结果。预测结果文件的格式与竞赛中的提交要求一致，结果文件请命名为 `result.csv`。初始的时候 `prediction_result/` 文件夹会被清空。

数据结构如下：

```
|-- prediction_result
    |-- result.csv
```

## 4、代码文件夹 `code/`

选手需要在此文件夹中存放预测与训练代码。代码审核会剔除有作弊行为、方案中无算法贡献的团队。参赛选手的模型要求可复现，复现偏差小于规定范围。

### a. 预测相关

- 请确保对**测试集**的预测结果可以由提交的代码产出，训练与预测流程所使用到的源码都要包含在提交的文件中，并在README.md文件中指导运行。
- 所使用的依赖（操作系统版本，MATLAB/Python的版本，需要安装的Python package，使用到的PyTorch，TensorFlow，MXNet的版本等）都需要在README文件中写明。
- 如果有需要编译的文件，请提供编译的脚本，并在 `README.md` 中说明编译所需的GCC、CUDA、CuDNN版本等依赖。
- 请提供 `test.sh` 文件作为程序入口，确保可以通过执行该文件来运行预测程序，得到最终结果，并将结果保存到上述的 `prediction_result/result.csv` 文件中。
- 读入文件的路径请使用相对路径，比如 `../xfdata/XX`

### b. 训练相关

- 请将训练相关代码也存放于此 `代码文件夹` 中，文件结构组织由选手自行决定，并在下述 `README.md` 文件中简略叙述训练流程并给出**训练脚本**。
- 请固定训练时的超参，例如学习率、batchsize等影响模型训练的参数。
- 如果对训练集进行了补充标注、数据增广等，下述 `README.md` 文件中说明，并给出增广代码与标注结果。

## 5、解决方案及算法介绍文件 `README.md`

- 请选手在 `README.md` 文件中介绍自己的解决方案及算法，并叙述模型训练复现流程。若选手提交的代码在运行时有需要特殊注意的内容，也请在该文件中一并说明
- 所使用的系统依赖（操作系统版本，Python/Matlab的版本）都需要在 `README.md` 文件中写明。如果使用Python环境，需提供**requirements.txt**以确保运行环境可复现 (Python package的版本依赖, 例如PyTorch, TensorFlow, MXNet、numpy等)。

## 6、附注

```
project
|--README.md           # 解决方案及算法介绍文件，必选
|--requirements.txt    # Python环境依赖
|--xfdata              # 比赛数据集
|--user_data           # 选手数据文件夹
    |--model_data      # 模型文件夹示例，可自行组织
        |--model.dat
    |--tmp_data         # 临时存储文件夹示例，可自行组织
        |--tmp.dat
|--prediction_result    # 预测结果
|--code                # 选手代码文件夹
    |--train           # 训练代码文件夹示例，可自行组织
    |--test            # 预测代码文件夹示例，可自行组织
    |--test.sh         # 预测执行脚本，必选
    |--train.sh        # 训练示例脚本，必选
```