**解释文档**

**队名: dm15331281 排名: 10**

**文件结构**：

ROOT

| code

| (注：此文件夹存放代码)

| | dm15331281\_10\_Code-feature.py

| | dm15331281\_10\_Code-model.py

| data

| (注：此文件夹存放数据和特征)

| | voice\_train

| | sms\_train

| | wa\_train

| | uid\_train

| | voice\_test\_b

| | sms\_ test\_b

| | wa\_ test\_b

| | train\_featureV1

| | test\_featureV1

| result

| (注：此文件夹存放预测结果)

| | lgb-baseline

**用到的库**：

Pandas 0.22.0，numpy 1.14.0，re，lightgbm，sklearn，matplotlib.pyplot，seaborn，xgboost

**环境**：win10，python3.6.4

**运行方法**：

1. 在code文件夹下用下面的命令提取特征：

python dm15331281\_10\_Code-feature.py

1. 在code文件夹下用下面的命令训练模型和预测：

python dm15331281\_10\_Code- model.py

1. 预测结果在result文件夹内

**特征**：

1. Voice特征：
2. Sms特征：
3. Wa特征：

**模型**：

由10个lgb和1个xgb融合，各自得出概率，再求概率均值。