所使用Python版本为3.6.1，使用的额外的库为scikit-learn、pandas、numpy、scipy、lightgbm。

建议使用软件anaconda，该软件集成了各种Python的科学计算扩展包，各个开源库安装便利。

流量部分整个工程的代码和数据的结构为：

/code/

--flowPreprocessing.py //预处理得到流量数据

--userPreprocessing.py //预处理得到用户数据

--flowPredictModelPart1.py //小区流量单点建模预测

--flowPredictModelPart2.py //小区流量24小时建模预测

--main.py //主函数文件用于运行整个代码模块，得到最终预测结果

/data/

--month5/data.csv //5月数据

--month6/data-1.csv //6月第一批数据

--month6/data-2.csv //6月第二批数据

/output/

--Common\_CellName\_flow\_data.csv //预处理得到的流量数据

--Common\_CellName\_new\_data.csv //预处理得到的用户数据

--modelResFir.csv //小区流量单点建模预测结果

--modelResSec.csv //小区流量24小时建模预测结果

--resultFinal.csv //小区流量最终建模预测结果

code文件夹存放源代码文件；

data文件夹为源数据文件；

output文件夹存储一些中间文件和最终生成结果；

运行前需要将数据放入data文件夹，将所给的excel文件转为csv文件，其中五月数据存放在data/month5文件夹目录下文件名为data.csv，六月数据存放在data/month6文件夹目录下文件名为data-1.csv，data-2.csv。

运行code文件夹中的main.py文件，该文件将依次运行code文件夹中的其他代码文件，

运行结束后output文件夹中将生成几个中间文件以及最终的结果文件resultFinal.csv。

Main文件运行逻辑顺序为：

with open('./flowPreprocessing.py','r') as f:

exec(f.read())

with open('./userPreprocessing.py','r') as f:

exec(f.read())

with open('./flowPredictModelPart1.py','r',encoding='UTF-8') as f:

exec(f.read())

with open('./flowPredictModelPart2.py','r',encoding='UTF-8') as f:

exec(f.read())

程序最后一次的控制台输出即为预测的准确率，其中mape损失误差为相对损失结果

结果文件resultFinal包含以下字段：CellName,Time,hour,pre,real，分别为小区ID，日期，小时，预测值，真实值。