1. **文件说明**

**evolution\_elm.m**; 主程序，用来训练和测试

**X\_1;** X的样本，每个城市4个数据，分别为 pm2.5,pm10,wind\_direction,wind\_speed,

共

**y\_1;** y的样本，每个城市一个值，分3类，取值为0,1,2,

**y\_2**; y的样本，每个城市一个值，取值为0,1

**gen\_data\_1.py;** 用来生成X\_1文件

**gen\_data\_y\_1.py**； 用来生成y\_1文件

**show\_city\_name.py;** 用来显示周围哪些城市的数据被用来预测

**aqi\_gt\_200;** 每个城市aqi大于200的天数统计，用来表征一个城市的污染严重程度

**city\_num\_X\_1**; X的样本中城市与序号的对应关系

**city\_num\_y\_1;** y的样本中城市与序号的对应关系

**dist\_matrix;** 城市之间距离的矩阵，用经纬度表示

**elm.m**; elm的代码

1. **实验结果**
2. **分两类时**

feature为所选城市的pm2\_5,pm10，wind\_direction,wind\_speed

选取最近的10个城市中污染最严重的8个城市，共210组数据，其中160组作为训

练数据，其余50组作为测试数据。

采用elm算法，节点数30

准确率 0.82

1. **分三类时**

feature为所选城市的pm2\_5,pm10，wind\_direction,wind\_speed

选取最近的10个城市中污染最严重的8个城市，共210组数据，其中160组作为训

练数据，其余50组作为测试数据。

采用elm算法，节点数25

准确率 0.60

补充：分两类时，随机猜测的准确率为0.5，我们的结果为0.8

分三类时，随机猜测的准确率为0.33，我们的结果为0.6

分三类以不同的城市作为目标差别会很大，因为有些城市label为2的情况很少，基本只有0和1，而北京的0，1，2分布较均匀，预测难度相对更高