1. 输入学生全省排名感觉不太合理，因为每个省份人数不一，同样的排名并不能有同样的地位。考虑改为排名百分比。即分数排名/全省高考人数

2. 大学数据中是大学在省份内的排名也不合理，比如云南省排名第一的学生肯定不会上云南排名第一的大学，大概率还是上全国排名第一的大学。所以大学数据中改为各个大学在全国的排名（在现有的大学中依次排）

所以对于每组来说：

1. 祁陶王高2邵组：生成数据具体为：(学生排名百分比, (学生经度,学生纬度), 大学数据)

其中大学数据和周吾君保持一致，导出为.csv，一条数据为一行

1. 周吾君组：生成数据具体为：( (大学经度,大学纬度), 大学在全国的排名)

经纬度不要用PCA做成一维。导出为.csv，一条大学数据为一行

1. 张妍谈笑组：计算H函数是在训练神经网络过程中进行的，因为要求H函数值越好。所以她们俩的工作内容要修改。

还有训练时可能存在的问题：

1. 对于经纬度，不同省份的差异并不是很大，比如天津和北京，差异只到了小数点后三位。不容易训练
2. 不同特征的量纲问题，数据规范化

对于经纬度问题：需要对经纬度按照一定比例进行缩放，周吾君负责的学校经纬度和大组负责的学生经纬度缩放标准要保持一致，缩放结果：还是两维数组，每一维的范围是[0,1]

周吾君：大学在全国的排名也要缩放到[0,1]之间