# 分工计划书

## 背景：新冠之下，为弘扬社会主义新时代的文化精神，贯彻东大人止于至善之高尚情操，本小组于今年上半年毅然投身于算法实现小组，立志于为每一位即将进入大学的莘莘学子提供引导和帮助。然而ddl将近而诸位繁忙，经综合权衡，并且在syz的push之下形成了此篇分工文档，还望大家海涵！

## 正文：

**周吾君**：

1、找出各个省份对应的经纬度

2、各个大学在所在省份内的排名

3、最后输出每个大学对应的编码：（经纬度+排名）【三维】或者对经纬度做pca成一维这样一共就是二维【如果出现大学空缺，则需要自动往前补位，以防止出现排名不连续的情况】

**胡艺凡**：

清洗数据模拟得到的数据，将各个大学的id对应成编码并且对接各个部分成员，将数筛选成三个dataset（文科、理科、综合）,使得数据集满足后续训练要求。

**徐浩卿&王倩**：

用神经网络训练文科、理科、综合三个模型

**张妍、谈笑**：PPT P7&P9

根据神经网络输出结果，计算H并且设计后续输出和筛选（功能扩展）。

**祁丁然&陶特&王昕彤&邵彤&高捷&邵一展**

每人分配若干所大学（第一组所有的大学数6等分），根据大学的ID、省份、文理科和分数，上下浮动地随机生成N个数据以供训练（N取值可以近似于一分一段表所对应的相邻两分数人数的差值。例如：河北省586分 生成多少人就可以找去河北省一分一段表中586和587差多少人得到N）【往上浮动的多，往下浮动的少】

##### 此分工仅为初代1.0版本，如有失偏颇之处，还望各位指正，对于分工任务有异议或者任何问题的，欢迎私戳我，阿里嘎多扩大伊马斯！

胡艺凡

2020/9/13于m3辅楼