重命名数据：

pre\_process.pretreatment.rename

可能会报错，一些数据无法重命名，需要手动更改

若数据都是以new开头，需要一下代码

pre\_process.pretreatment.rename\_del\_new

检查数据：

pre\_process.pretreatment.check\_dir

提取特征：

pre\_process.feature\_extractor.extract\_feature

三期分别生成csv格式

将三期特征合并到一个文件：

pre\_process.pretreatment.merge\_3\_period

得到all\_features.csv

计算ICC：

根据一个group中的case名，从group1中获得相应的case：

process.compute\_ICC.pick\_group1\_data

需要在D:\PycharmProjects\learning\SHZL\feature\_data\维度\新的ICC\创建相应表格

pre\_process.feature\_extractor.extract\_feature提取特征

三个文件及分别 pre\_process.merge\_3\_period得到all\_features.csv

获得ICC值表格：

process.pretreatment.compute\_ICC.compute\_all\_ICC

将各个ICC表格进行汇总至ICC.csv

process.feature\_selector.merge\_ICC

根据ICC筛选特征：

查看ICC分布直方图

process. feature\_selector. show\_ICC\_distribution

可以看一下ICC值的百分位数

process feature\_selector.compute\_percentile

计算ICC低于阈值的特征个数

process. feature\_selector.compute\_num\_ender\_threshold

算算相关性？？？虽然不知道有啥用

process. feature\_selector.compute\_corr

通过阈值筛选特征，输入的是维度文件夹，得到selected\_features.csv

process. feature\_selector.select\_feature\_by\_threshold

模型数据准备：

根据selected\_feature.csv，将数据分到fuhrman和isup文件夹中，并加入label

（会有一些红字，施主莫要惊慌）

process. feature\_selector.seperate\_furman\_isup

*将数据以7:3的比例拆分成训练集和测试集*

*Process.seperate\_features.seperate\_train\_test*

*根据 standard\_feature.csv 将train和test数据分配到三期中*

*Process.seperate\_features. seperate\_traintest\_period\_data*

*对三个训练集文件做Smote*

*将不同类特征进行分配，需要实现创建几个 “n”kind\_feature 文件夹*

*process.seperate\_features. seperate\_3period\_features*

先将比如ISUP\_feature.csv的分配到各个period，再分配到nkind文件夹里

（代码遗失，应该是根据Process.seperate\_features. seperate\_traintest\_period\_data改的）

用Fae拆分（根据test）

根据fae的拆分方式，拆分原始standard\_features.csv文件

Process.seperate\_features.seperate\_train\_test

手动把生成的train\_data和test\_data复制到相应data文件夹里

慢慢跑一大堆模型，然后挑一些结果好的出来，共9个

联合模型：

挑选模型

Process.select\_model.copy\_model

训练三期联合模型

Process.united\_model.unite\_3model

可以根据输出的prob值来检验模型之间是否存在显著差异

Process.examine\_models.wilcoxon\_test

也可以用bootstrap方法来计算模型的auc值置信区间

Process.examine\_models.compute\_ic（联合模型）

Process.examine\_models.compute\_period\_ic（期态模型）