

Problem 1

分类问题：病症分类

- (1) 目标变量是病症类型，特征有病人症状，检查结果，病人年龄、病史等匿名资料数据。
- (2) 可在公开疾病数据库收集数据，如 39 疾病百科等
- (3) 并未全面解决，目前主要是生物医学图像处理辅助判断疾病。

Problem 2

```
from numpy import *  
n,p=[200,20];  
  
r_feature_mat=mat(random.rand(n,p));  
t_variable_vec=arange(n);  
  
train=random.choice(n,int(n*0.7),replace=False);  
training_dataset=[r_feature_mat[item] for item in train];  
  
test=set(arange(n))-set(train);  
test_dataset=[r_feature_mat[item] for item in test];  
  
save("train_data.npy",training_dataset);  
save("test_data.npy",test_dataset);
```