

### Problem 1

垃圾邮件的过滤

- 1、目标变量是邮件，垃圾邮件的潜在特征是发送者未知，未经许可强制发送，信件开头有“取消订阅”等字样。
- 2、可以在自己的邮箱内寻找，可以在网上寻找垃圾邮件的模板、样本。
- 3、对于一般的垃圾邮件可以过滤，但少部分“逼真”的垃圾邮件无法过滤。提取邮件的特征，做二分类，看属于垃圾邮件还是正常邮件。

### Problem 2

```
import numpy as np
n=200
p=20
matrix=np.random.rand(n,p)
vector=np.random.rand(n,1)
row_rand_array = np.arange(matrix.shape[0])
np.random.shuffle(row_rand_array)
train=matrix[row_rand_array[0:140]]
text=matrix[row_rand_array[140:n]]
np.savetxt("train.txt",train)
np.savetxt("text.txt",text)
```