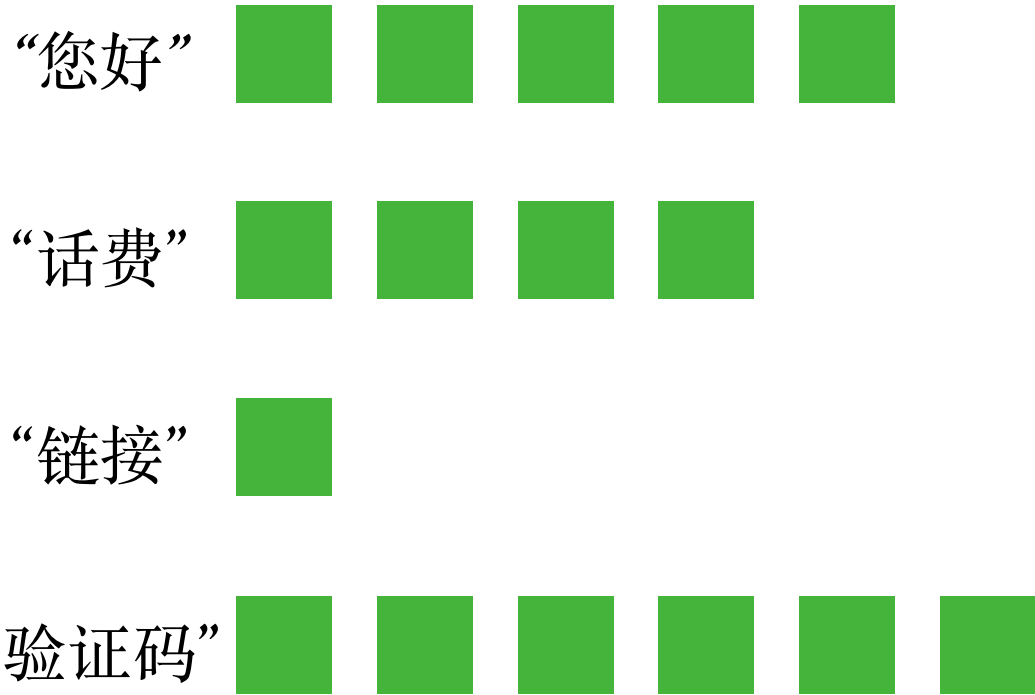
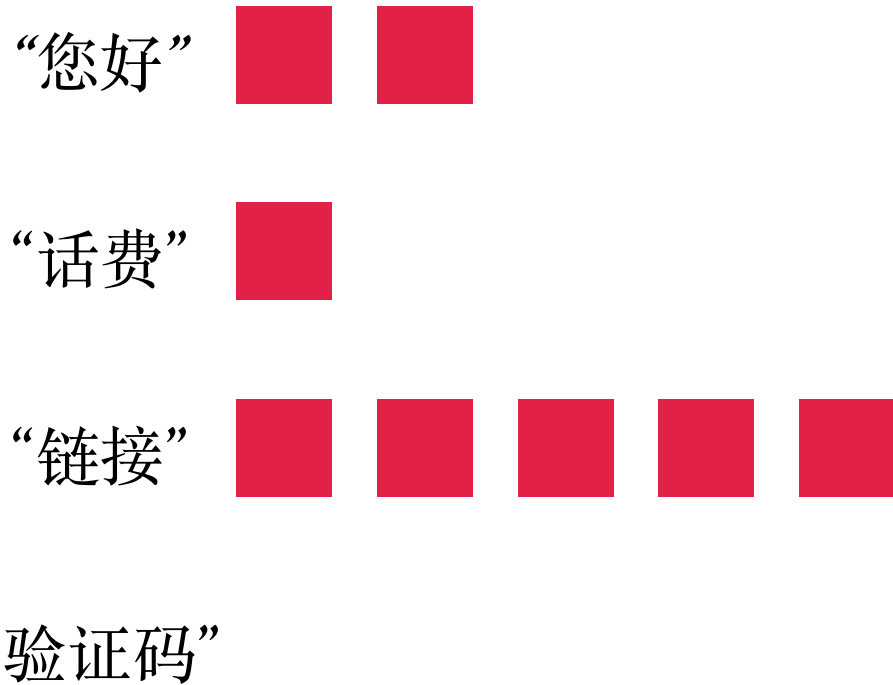


1. 计算机-分类:

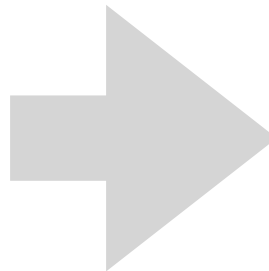
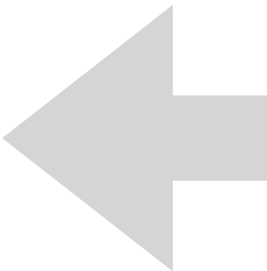
正常短信
9条



垃圾短信
3条



首先，计算机对两类短信中的特定词语进行统计，
计算出**各个词语**在两类中出现的**次数**



2. 计算机-计算概率：

正常短信
9条



“您好” 

$$P(\text{“您好”}|\text{正常}) = \frac{5}{16}$$

“话费” 

$$P(\text{“话费”}|\text{正常}) = \frac{4}{16}$$

“链接” 

$$P(\text{“链接”}|\text{正常}) = \frac{1}{16}$$

“验证码” 

$$P(\text{“验证码”}|\text{正常}) = \frac{6}{16}$$

其次，两类短信中的特定词语进行条件概率计算，

$$= \frac{\text{该词在此类中出现的次数}}{\text{此类中的总词数}}$$

垃圾短信
3条



“您好” 

$$P(\text{“您好”}|\text{垃圾}) = \frac{2}{8}$$

“话费” 

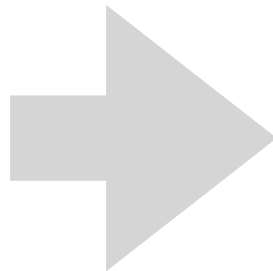
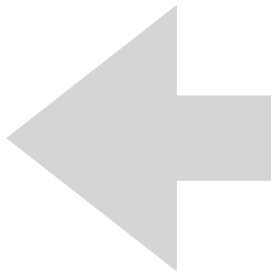
$$P(\text{“话费”}|\text{垃圾}) = \frac{1}{8}$$

“链接” 

$$P(\text{“链接”}|\text{垃圾}) = \frac{5}{8}$$

“验证码”

$$P(\text{“验证码”}|\text{垃圾}) = \frac{0}{8}$$



情景：



在过去一段时间，你手机上一共收到了12条短信。
经过你的判断，发现其中9条为**正常短信**，如话费短信、验证码等。
但其余3条都为**垃圾短信**，如广告。

现在你希望计算机能够对这些已经收到的短信进行学习，
并且帮你**判断**之后收到的短信是**正常短信**还是**垃圾短信**。

