

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”

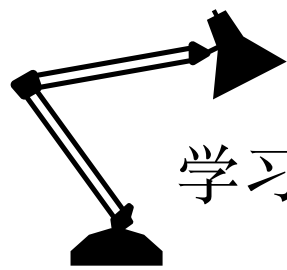


$$P(\text{“链接”}|\text{垃圾}) = \frac{5}{8}$$

“验证码”

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

2. 计算机-应用结果：



学习成果：

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

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

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

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

正常短信
9条



至此，计算机已经差不多完成了对已收到的短信的学习。

接下来就可以让计算机帮助我们判断新收到短信的类型，来验证它学习的成果。

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

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

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

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

垃圾短信
3条

