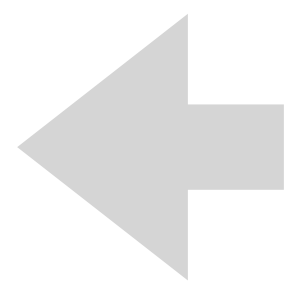


Made by Mike_Zhang
<https://ultrafish.cn>




4. 总结

筛选

学习

应用




正常短信
9条
 $P(\text{正常}) = 0.75$

“您好”

“话费”

“链接”

“验证码”



垃圾短信
3条
 $P(\text{垃圾}) = 0.25$

“您好”

“话费”

“链接”

“验证码”

$$P(\text{“您好”}|\text{正常}) = \frac{5 + 1}{16 + 4} = \frac{6}{20}$$
$$P(\text{“话费”}|\text{正常}) = \frac{4 + 1}{16 + 4} = \frac{5}{20}$$
$$P(\text{“链接”}|\text{正常}) = \frac{1 + 1}{16 + 4} = \frac{2}{20}$$
$$P(\text{“验证码”}|\text{正常}) = \frac{6 + 1}{16 + 4} = \frac{7}{20}$$

$$P(\text{“您好”}|\text{垃圾}) = \frac{2 + 1}{8 + 4} = \frac{3}{12}$$
$$P(\text{“话费”}|\text{垃圾}) = \frac{1 + 1}{8 + 4} = \frac{2}{12}$$
$$P(\text{“链接”}|\text{垃圾}) = \frac{5 + 1}{8 + 4} = \frac{6}{12}$$
$$P(\text{“验证码”}|\text{垃圾}) = \frac{0 + 1}{8 + 4} = \frac{1}{12}$$

“验证码 链接 链接 链接”



$$P(\text{正常}) \times P(\text{“验证码”}|\text{正常}) \times P(\text{“链接”}|\text{正常})^3 = 0.75 \times \frac{7}{20} \times (\frac{2}{20})^3 = 0.0002625$$

“验证码 链接 链接 链接” 为正常短信的概率

$$P(\text{垃圾}) \times P(\text{“验证码”}|\text{垃圾}) \times P(\text{“链接”}|\text{垃圾})^3 = 0.25 \times \frac{1}{12} \times (\frac{6}{12})^3 = 0.0026$$

“验证码 链接 链接 链接” 为垃圾短信的概率

0.0026 > 0.0002625, 垃圾短信。