4. 总结



正常短信

9条

P(正常) = 0.75

"链接"

"验证码"

"您好"



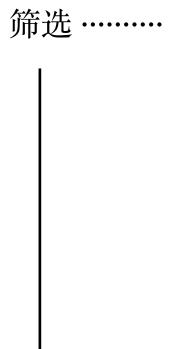
垃圾短信

3条

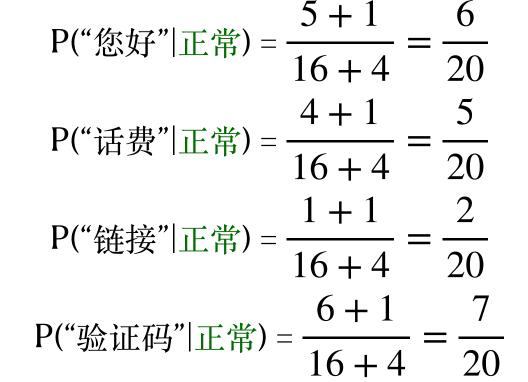


P(垃圾) = 0.25

"验证码"



学习



 $P(\text{"$xy"}|\frac{1}{2}) = \frac{2+1}{8+4} = \frac{3}{12}$ $P("话费"|<mark>垃圾) = \frac{1+1}{8+4} = \frac{2}{12}$ </mark> $P(\text{"链接"}|\text{垃圾}) = \frac{5+1}{8+4} = \frac{6}{12}$ P("验证码"|垃圾) = $\frac{0+1}{8+4} = \frac{1}{12}$

"验证码链接链接链接"

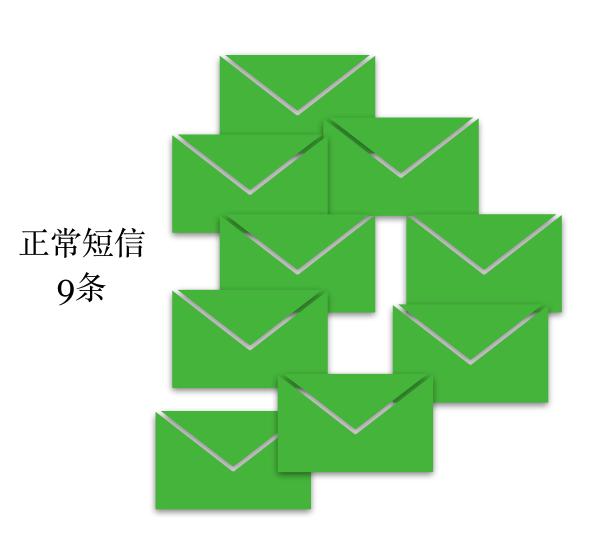


 $P(正常) \times P("验证码"|正常) \times P("链接"|正常)^3 = 0.75 \times \frac{7}{20} \times (\frac{2}{20})^3 = 0.0002625$ "验证码链接链接链接"为正常短信的概率

 $P(垃圾) \times P("验证码"|垃圾) \times P("链接"|垃圾)³ = 0.25 × <math>\frac{1}{12} \times (\frac{6}{12})^3 = 0.0026$

"验证码链接链接链接"为垃圾短信的概率

3. 计算机-学习优化:



P(正常) = 0.75

垃圾短信 3条

P(垃圾) = 0.25

$$P("您好"|正常) = \frac{5+1}{16+4} = \frac{6}{20}$$

$$P("话费"|正常) = \frac{4+1}{16+4} = \frac{5}{20}$$

$$P("链接"|正常) = \frac{1+1}{16+4} = \frac{2}{20}$$

$$P("验证码"|正常) = \frac{6+1}{16+4} = \frac{7}{20}$$

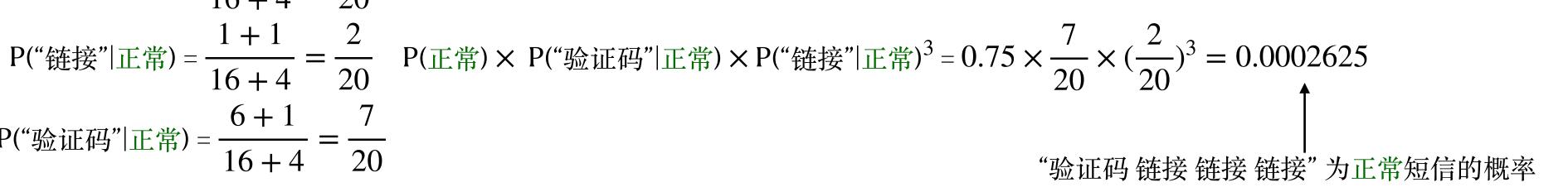
$$P("您好"|垃圾) = \frac{2+1}{8+4} = \frac{3}{12}$$

$$P("话费"|垃圾) = \frac{1+1}{8+4} = \frac{2}{12}$$

$$P("链接"|垃圾) = \frac{5+1}{8+4} = \frac{6}{12}$$

$$P("验证码"|垃圾) = \frac{0+1}{8+4} = \frac{1}{12}$$

此时,我们手机又收到一条短信: "验证码链接链接链接"



再把优化后的学习结果应用到此短信上,

0.0026 > 0.0002625

计算机得出此短信为一条垃圾短信,符合实际。

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

"验证码链接链接链接"为垃圾短信的概率

