

Осенов, ip. 492

## Трёхстишие загадки

14.2

4.2 ROC-AUC (треугольник)

Пусть  $P(a(x)=1)=p$ ,  $P(a(x)=0)=1-p$ , тогда  
 $TPR = \frac{\text{True pos}}{\text{True pos} + \text{false neg}} = \frac{px}{px + (1-p)x} = \frac{p}{1}$  если  $x$  элементов из класса 1

$$FPR = \frac{\text{False pos}}{\text{False pos} + \text{True neg}} = \frac{p(1-x)}{p(1-x) + (1-p)(1-x)} = p$$

T.K.  $x \in [0, 1]$   $\diamond$

$$p \in [0, 1] \Rightarrow$$

кривая ROC-AUC - это кривая  $y=x \Rightarrow$

Значение в среднем  $S_{\text{удел}} = 0,5$

4.1

#### 4.1 Наивный байесовский классификатор

$$P(y|x) = \frac{P(x|y) P(y)}{P(x)}$$

$$\hat{a}(x) = \operatorname{argmax}_y P(y|x)$$

В предположении, что  $P(y)$  равны, а  $P(x^{(i)}, y)$  независимы получаем

$$a(x) = \operatorname{argmax}_y P(y) \prod_{k=1}^n P(x^{(k)} | y)$$

$$L(x) = \ln a(x) \propto \operatorname{argmax}_y \left[ \ln P(y) - \sum (\mu_{y_k} - x^{(k)})^2 \right] \rightarrow$$



или  $20,17 \rightarrow P \in [0,1] \rightarrow$

кривая ROC-AUC - это площадь  
 $y=x \rightarrow$   
Значение в среднем будет 0,5

4.1 Наивный байесовский классификатор.

$$P(y|x) = \frac{P(x|y) P(y)}{P(x)}$$

$$\hat{a}(x) = \arg \max_y P(y|x)$$

В предположении, что  $P(y)$  равны, а разл.  $P(x^{(k)}|y)$  независимы получаем

$$\hat{a}(x) = \arg \max_y P(y) \prod_{k=1}^K P(x^{(k)}|y)$$

$$L(x) = \ln \hat{a}(x) \approx \arg \max_y \left[ \ln P(y) + \sum (\mu_{y_k} - x^{(k)})^2 \right] \Rightarrow$$

~~Если~~  $\mu_{y_k}$  находимые близки к  $x$ , то  
 $L(x)$  будет ~~находиться там~~.