Милогина Александра, 496

3.1. Знакомство с линейными классификаторами

1. Бинарный линейный классификатор

$$a(x) = sign(\langle w, x \rangle + w_0)$$

2. Отступом алгоритма $a(x) = sign(f(x))$ на объекте x_i называется величина $M_i = y_i f(x_i)$

(y_i – класс, к которому относится x_i)

Знак отступа:

$M_i \leq 0 \Leftrightarrow y_i \neq a(x_i)$

$M_i > 0 \Leftrightarrow y_i = a(x_i)$

3. Классификаторы вида $a(x) = sign(\langle w, x \rangle - w_0)$ сводятся к классификаторам вида $a(x) = sign(\langle w, x \rangle)$ путем введения фиктивного константного признака $x_0 = -1$

4. Функционал эмпирического риска:

$$Q(w) = \sum_{i=1}^{l} [M_i(w) < 0]$$

Для наилучшего алгоритма классификации: $Q(w) = 0$

5.

$M_i < 0$

$a(x) = sign(\langle w, x \rangle)$

w?

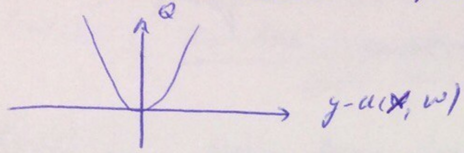
Параметр $w^* = 0$ минимизирует функционал

6. Функ-л аппроксимированного эмпирического риска

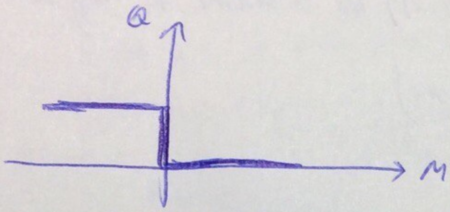
$$\tilde{Q}(w) = \sum_{i=1}^{l} L(M_i(w)),$$ где $M_i(w)$ – отступ на i-ом объекте

7. Функция потерь $Q(w)$ отражает ошибку алгоритма $a(x, w)$ на выборке X. Минимизация ф-ции потерь приводит к построению алгоритма

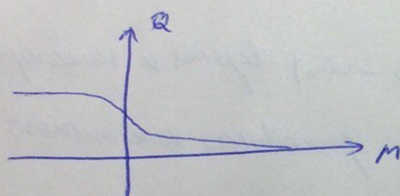
который ~~...~~ оптимально угадывает ответы $\hat{y}$ на выборке X для функц. $Q(w)$.



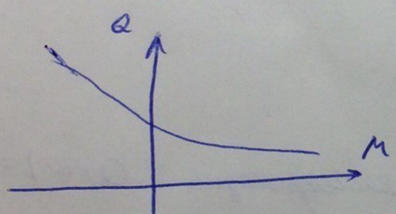
"квадратичная"



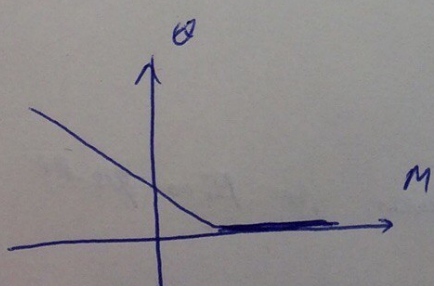
пороговая



сигмоидная



логистическая



кусочно-гладкая

8. Пример негладкой ф-ции потерь:

$$Q(w) = |y - a(x, w)|$$

15. Accuracy – доля правильных ответов для классификации

$$\text{Precision} = \frac{\text{число верно классифицированных объектов целевого класса}}{\text{общее число объектов}}$$

$$\text{Recall} = \frac{\text{число верно классифицированных объектов целевого класса}}{\text{общее число объектов целевого класса}}$$