ML THW 1 Yaropolov MAE 21

Задача 2

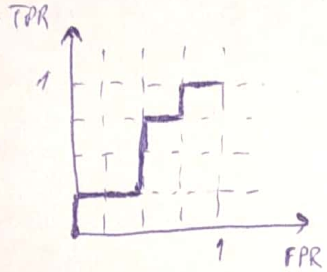
Упорядочим:

i	$b(x_i)$	y_i
8	.95	1
5	.9	-1
2	.8	-1
7	.6	1
6	.3	1
4	.25	-1
3	.2	1
1	.1	-1

$\Rightarrow$ для соотв. порогов

TPR	FPR
1/4	0
1/4	1/4
1/4	2/4
2/4	2/4
3/4	2/4
3/4	3/4
1	3/4
1	1
$\frac{TP}{TP+FN}$	$\frac{FP}{FP+TN}$

$\Rightarrow$



- строим AUC кривую по вычисленным TPR, FPR
- AUC-ROC = площадь под кривой $= \frac{9}{9+7} = \frac{9}{16}$
- пороги t задаем так, чтобы отделять по одному наблюдению

Задача 3

Рассмотрим посл. $y_{(1)}, \dots, y_{(\ell)}$. Она задает кусочно-ломанную AUC, состоящую из $+1 \equiv |$, $-1 \equiv -$. Рассмотрим случайный объект $y_{(i)} = -1$. $P(b(x_{(j)}) \geq b(x_{(i)}) \mid y_{(j)} = 1, y_{(i)} = -1) = \frac{\sum_{j=i}^{\ell} [y_{(j)} = 1]}{\ell_+}$ (1), т.е. ~~кол-во~~ доля положительных объектов, идущих за $y_{(i)}$. С другой стороны, $y_{(i)} = -1$ соответствует горизонтальный отрезок AUC, площадь под которым равна $\frac{1}{\ell_-} \cdot \frac{\sum_{j=i}^{\ell} [y_{(j)} = +1]}{\ell_+}$. Чтобы получить AUC-ROC, надо сложить эти маленькие площади по всем i, где $y_{(i)} = -1$ $\Rightarrow$ AUC-ROC $= \sum_{i=1}^{\ell} \frac{[y_{(i)} = -1]}{\ell_-} \cdot \sum_{j=i}^{\ell} \frac{[y_{(j)} = 1]}{\ell_+}$. Но это и есть вероятность (1), просуммированная с весами $\frac{[y_{(i)} = -1]}{\ell_-}$ - по всем i, т.е. с весами - вероятностями того, что случайно выбранный объект $= -1$. Т.е. выбирая случайно i, и к нему j (такие что $y_{(i)} = -1$, $y_j = 1$) мы получим сумму вероятностей, равную AUC-ROC.

Задача 4

$$\begin{cases} \frac{1}{2}\|w\|^2 \to \min\limits_{w,b,\xi} \\ y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i \quad (1) \\ \xi_i \geq 0 \end{cases}$$

Пусть $w = 0$, тогда $\forall b$ - для любого найдутся такие большие положительные ξ_i, что (1) будет выполнено. Уменьшить $\|w\|^2$ больше нельзя $\Rightarrow$ это решение. (Можно взять в точности $\xi_i = |1 - y_i(\langle w, x_i \rangle + b)|$)

Задача 5

- $\sigma'(z) = -\frac{1}{(1+\exp(-z))^2} \cdot \exp(-z) \cdot (-1) = \frac{1}{1+\exp(-z)} \cdot \frac{\exp(-z)}{1+\exp(-z)} = \sigma(z) \cdot (1 - \sigma(z))$

- $L(x, y; w) = \ln(1 + \exp(-y\langle w, x \rangle)) = -\ln \sigma(-y\langle w, x \rangle) \Rightarrow \nabla_w L(x, y; w) = -\frac{1}{\sigma(-y\langle w, x \rangle)} \cdot \sigma(-y\langle w, x \rangle)(1 - \sigma(-y\langle w, x \rangle)) \cdot (-y \cdot x) = [\sigma(-y\langle w, x \rangle) - 1] \cdot y \cdot x$

Задача 1

$Q(w^{k-1} - \alpha \nabla_w Q(w^{k-1})) \to \min\limits_{\alpha} \Rightarrow$ FOC wrt α $\langle \nabla_w Q(w^{k-1} - \alpha \nabla_w Q(w^{k-1})), -\nabla_w Q(w^{k-1}) \rangle = 0$

Обозначим $\nabla_w Q(w^{k-1})$ как q. Заметим $\nabla_w Q(w) = \frac{2}{\ell} \sum (\langle x_i, w \rangle - y_i) \cdot x_i + 2\lambda w$. Тогда:

$$\langle \frac{2}{\ell} \sum (\langle x_i, w^{k-1} - \alpha q \rangle - y_i) x_i + 2\lambda (w^{k-1} - \alpha q), q \rangle = 0$$

$$\langle \frac{1}{\ell} \sum (\langle x_i, w^{k-1} \rangle - y_i) x_i - \alpha \cdot \frac{1}{\ell} \sum \langle x_i, q \rangle x_i + \lambda w^{k-1} - \alpha \lambda q, q \rangle = 0$$

$$\alpha = \frac{\langle \frac{1}{\ell} \sum (\langle x_i, w^{k-1} \rangle - y_i) x_i + \lambda w^{k-1}, q \rangle}{\langle \frac{1}{\ell} \sum \langle x_i, q \rangle x_i + \lambda q, q \rangle}$$

- ответ, наилучшая длина шага.