Обучение ранжированию

Методы: pointwise
pairwise
✓ listwise

Списочные методы

ListNet

истинная релевантность документа на i-й позиции в ответ на запрос q

$$DCG@k = \frac{1}{|Q|} \sum_{q \in Q} \sum_{i=1}^{k} \frac{2^{y(i)}}{\log(i+1)}$$

$\frac{\partial}{\partial w} DCG@k$ - ?

$a(q,d): \; a(q, d_{(1)}) \geq a(q, d_{(2)}) \geq \dots$

Заменим DCG@k(q) на:

$$\sum_{\pi} p(\pi) \cdot DCG@k(\pi)$$

сумма по всем перестановкам документов

вероятность, дифференцируемая по модели

3 документа: d_1, d_2, d_3

$d_1\, d_2\, d_3$	$p(\;)$
$d_1\, d_3\, d_2$	$p(\;)$
$d_3\, d_1\, d_2$	$p(\;)$
$d_3\, d_2\, d_1$	⋮
$d_2\, d_1\, d_3$	⋮
$d_2\, d_3\, d_1$	

$a(d_2) > a(d_1) > a(d_3)$

Рассматриваем запрос q.

Нужно отранжировать $d_1, \dots d_{n_q}$

$y_1, \dots y_{n_q}$ - истинные оценки релевантности

$z_1, \dots z_{n_q}$ - оценки релевантности у модели

$| \; a(d_1) \quad \ldots \quad a(d_{n_q})$

Введем распределение на всех перестановках на n_q документах

$$P_z(\pi) = \prod_{j=1}^{n_q} \frac{\varphi(z_{\pi(j)})}{\sum_{k=j}^{n_q} \varphi(z_{\pi(k)})}$$, φ - любая неубывающая строго положит. функция

$$\varphi(z) = e^z$$

Св-ва: 1) $P_z(\pi)$ задает распределение на S_{n_q}

2) π: d_i выше d_j , $z_i > z_j$

π': как π, только $d_i \leftrightarrow d_j$ (т.е. π' хуже, чем π)

$P_z(\pi') < P_z(\pi)$

3) макс. вероятность имеет сортировка документов по убыванию z_i

мин. вероятность – по возрастанию z_i

$$F(a) = \frac{1}{|Q|} \sum_{q \in Q} \sum_{\pi \in S_{n_q}} \underbrace{P_z(\pi)}_{= \prod_{j=1}^{n_q} \frac{\exp(a(q, d_{(j)}))}{\sum_{k=j}^{n_q} \exp(a(q, d_{(k)}))}} \cdot DCG@k(\pi) \to \max$$

$\frac{\partial F}{\partial w}$ существует

проблема: $n_q!$ перестановок

<u>Вариант 1.</u> Сгенерировать m перестановок из $P_z(\pi)$, суммировать только по ним

<u>Вариант 2.</u> Перейти от $P_z(\pi)$ к $P_z(j)$

вероятность

попадания d_j
на первую позицию

$$P_z(j) = \frac{\varphi(z_j)}{\sum_{k=1}^{n_q} \varphi(z_k)}$$

$$z_i > z_j \Rightarrow P_z(i) > P_z(j)$$

$$-\frac{1}{|Q|} \sum_{q \in Q} \underbrace{\sum_{j=1}^{n_q} P_y(j) \log P_z(j)}_{\text{кросс-энтропия}}$$ — оптимизируем

Обобщенные линейные модели (GLM)

Линейная регрессия

$$\frac{1}{c} \exp\left(-\frac{(y_i - \langle w, x_i \rangle)^2}{2\sigma^2}\right)$$

$$y \sim \mathcal{N}(\langle w, x \rangle, \sigma^2)$$

$$\log p(Y \mid X) = -\sum_{i=1}^{\ell} (\langle w, x_i \rangle - y_i)^2 \to \max$$

Логистическая регрессия

$$y \sim \text{Bernoulli}(\mu(x))$$

$$p(y \mid \mu(x)) = \mu(x)^y (1-\mu(x))^{1-y}$$

$$\mu(x) \in [0, 1]$$

$$\mu(x) = \sigma(\langle w, x \rangle)$$

---

y — счетчик событий

$$y \sim \text{Poisson}(\mu(x))$$

$$p(y \mid \mu(x)) = e^{-\mu(x)} \frac{\mu(x)^y}{y!}$$

$$\mu(x) > 0$$

$$\mu(x) = e^{\langle w, x \rangle}$$

Предположим, что плотность y — из экспоненц. семейства распределений

$$p(y \mid \eta) = \frac{h(y)}{g(\eta)} \exp(\langle \eta, s(y) \rangle)$$

обычно: $s(y) = y$

параметры

$\eta \in \mathbb{R} \Rightarrow \boxed{\eta(x) = \langle w, x \rangle}$ – $\langle w, x\rangle$ предсказывает значение параметра

$$\text{Обучение:} \quad \sum_{i=1}^{l} \log p(y_i \mid \underbrace{\eta(x_i)}_{=\langle w, x_i \rangle}) \to \max_w$$

$a(x) = ?$

$$a(x) \approx \mathbb{E}(y \mid x) = \int p(y \mid \eta(x)) \cdot y \, dy$$

как найти явную формулу для $a(x)$?

$$\frac{\partial \ln g(\eta)}{\partial \eta_j} = \frac{1}{g(\eta)} \cdot \frac{\partial}{\partial \eta_j} g(\eta) =$$

$g(\eta)$ – нормировочная константа, т.е. $\frac{1}{g(\eta)} \cdot \int h(y) \exp(\langle \eta, s(y) \rangle) dy$

$$= \frac{1}{g(\eta)} \frac{\partial}{\partial \eta_j} \int h(y) \exp(\langle \eta, s(y) \rangle) \, dy =$$

$$= \frac{1}{g(\eta)} \int h(y) \boxed{s_j(y)} \exp(\langle \eta, s(y) \rangle) \, dy =$$

$$= \int s_j(y) \, p(y \mid \eta) \, dy = \mathbb{E} s_j(y)$$

$$\nabla_\eta \ln g(\eta) = \mathbb{E} s(y) = \mathbb{E}(y \mid \eta) = a(x)$$

обычно $s(y) = y$

Алгоритм применения GLM:

1) Задаем распред. из экспон. семейства:

$p(y \mid \eta(x))$

2) $a(x) = \psi(\langle w, x \rangle) : \nabla_z \ln g(z) \Big|_{z = \langle w, x \rangle}$

3) ищем w через макс. правдоподобия