Учет предпочтения экспертов в задаче ранжирования.

Алсаханова Надежда

Москва, 2021

Постановка задачи

Задача

Получить рейтинг продуктов, учитывающий экспертные оценки и показатели продукта, с учетом наличия предпочтения на множестве экспертов.

Используемый метод:

Линейное согласование.

Код:

https://github.com/NadezhdaAlsahanova/Ratings.

Описание данных

Описания объектов (продуктов питания) представлено в виде матрицы исходных данных:

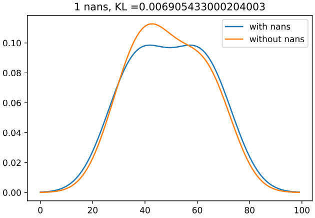
$$A = \{a_{i,j}\}_{i,j=1}^{m,n}$$

где m – количество объектов, n – количество показателей.

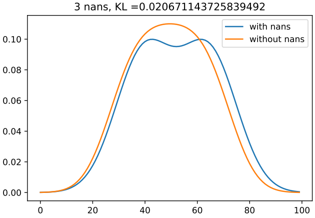
Каждый эксперт дает оценку весовых коэффициентов $\mathrm{w}_{0,k}$ и оценку объектов $\mathrm{q}_{0,k}$, где k - номер эксперта.

Критерий отказа от эксперта:

Изменение распределения оценок после заполнения пропусков предлагается измерять по дивергенции Кульбака-Лейблера.
На рисунке 1 приведены примеры изменения дивергенции Кульбака-Лейблера от количества пропусков. Для этого добавлялся эксперт со случайными оценками с различным количеством пропусков, затем пропуски заполнялись, как описано выше.



(a)



(b)

Рис.: Распределения до и после заполнения пропусков.

Линейное согласование

Экспертные оценки несогласованны, то есть $q_0 \neq Aw_0$. Необходимо провести согласование. Для этого воспользуемся линейным согласованием:

$$q_\alpha = \alpha q_0 + (1-\alpha)Aw_0$$

$$w_\alpha = (1-\alpha)w_0 + \alpha A^+ q_0$$

Где $\alpha \in [0,1]$ - параметр доверия экспертным оценкам объектов, либо весов. При $\alpha = 0$ мы игнорируем экспертные оценки объектов, при $\alpha = 1$ мы игнорируем экспертные оценки весов.

Оптимизационная задача:

$$\alpha^* = arg\min_\alpha \{\frac{1}{n}\|w_\alpha - w_0\|_2^2 + \frac{1}{m}\|q_\alpha - q_0\|_2^2\}$$

Линейное согласование

Чтобы проводить линейное согласование нужно перевести в вещественной пространство матрицу Q. Для этого предлагается использовать 'простые' рейтинги, полученные из вещественных признаков продуктов Pr, и корреляцию Кенделла:

$$\tau = 1 - \frac{4}{n(n-1)} R$$

где

$$R = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} [[x_i \; < \; x_j] \neq [y_i \; < \; y_j]]$$

Получаем вещественные признаки из оценок экспертов с помощью формулы:

$$\hat{Q}_i = \sum_{j=1}^{p} \tau(Pr_j, Q_i) * P_j$$

где p - количество вещественных признаков, P_j - столбец значений вещественного признака, а Pr_j - его 'простой' рейтинг.

Предподчтение экспертов

Пусть у нас есть рейтинг экспертов $\vec{r} \in \mathbb{R}^{\mathbb{K}}$, где K - количество экспертов., состоящий из неповторяющихся оценок от 1 до K. Чем выше оценка эксперта, тем предпочтительнее его оценки.

В качестве q_0 и w_0 были взяты взвешанные средние значения по $\mathrm{q}_{0,\mathrm{k}}$ и $\mathrm{w}_{0,\mathrm{k}}$.

Результаты эксперимента

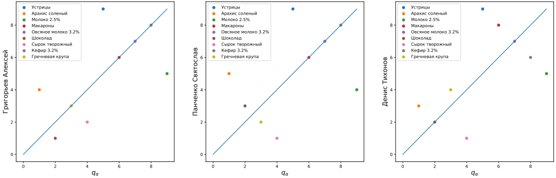


Рис.: Без учета рейтинга экспертов

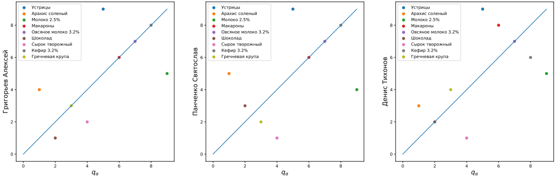


Рис.: С учетом рейтинга экспертов