Учет предпочтения экспертов в задаче ранжирования.

Алсаханова Н. Ю.

1 Постановка задачи

Заданы описания объектов (продуктов питания) 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 $\mathbf{w}_{0,k} \in \mathbb{R}^n$ и оценку объектов $\mathbf{q}_{0,k} \in \mathbb{R}^m$, где k - номер эксперта. Необходимо построить согласованный рейтинг объектов, то есть найти $\mathbf{q}_\alpha$ и $\mathbf{w}_\alpha$ такие, что:

$$\begin{aligned} \mathbf{q}_\alpha &= A\mathbf{w}_\alpha \\ \mathbf{w}_\alpha &= A^+\mathbf{q}_\alpha \end{aligned} \tag{1}$$

Для построения согласованного рейтинга применяется линейное согласование:

$$\mathbf{q}_\alpha = \alpha\mathbf{q}_0 + (1-\alpha)A\mathbf{w}_0$$

$$\mathbf{w}_\alpha = (1-\alpha)\mathbf{w}_0 + \alpha A^+\mathbf{q}_0$$

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.

Если α не задано экспертами, то его можно найти путем решения оптимизационной задачи:

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

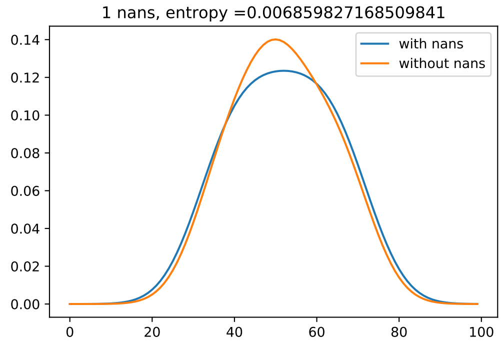
2 Эксперимент

2.1 Данные и заполнение пропусков

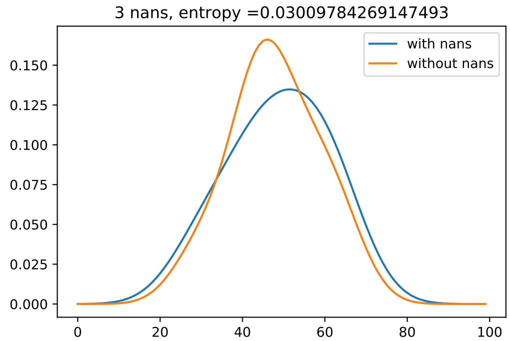
Данные для эксперимента состояли из описания (5 признаков) и K оценок для 9 объектов. Оценки рассматривались, как признаки, то есть матрица $A \in \mathbb{R}^{9\times K+5}$. Так же были даны оценки экспертов для каждого признака, $W \in \mathbb{R}^{17\times K'}$. Множество экспертов, оценивших объекты, не совпадало с множеством экспертов, оценивших признаки. Так же в матрицах A и W присутствовали пропуски. Для заполнения пропусков использовались средние значения 2 ближайших соседей.

Критерий отказа от эксперта: Если в оценках эксперта имеется слишком много пропусков, так что заполнение по ближайшим соседям сильно меняет распределение оценок, то предлагается не учитывать этого эксперта. Изменение ряспределение предлагается измерять по дивергенции Кульбака-Лейблера.

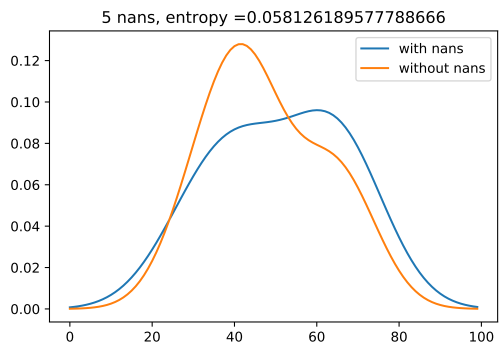
На рисунке 1 приведены примеры изменения дивергенции Кульбака-Лейблера от количества пропусков. Для этого добавлялся эксперт со случайными оценками с различным количеством пропусков, затем пропуски заполнялись, как описано выше.



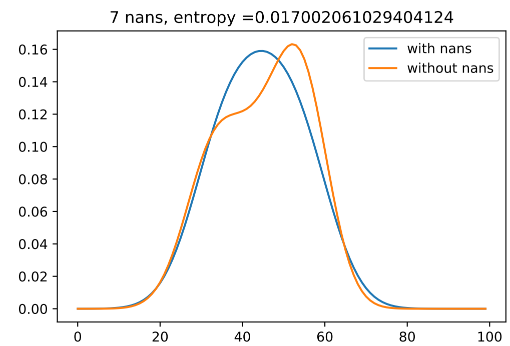
(а)



(б)



(в)



(г)

Рис. 1 Распределения до и после заполнения пропусков.

2.2 Согласование

В качестве $\mathbf{q_0}$ и $\mathbf{w_0}$ были взяты медианные значения по $\mathbf{q_{0,k}}$ и $\mathbf{w_{0,k}}$. Затем было проведено согласование. Результаты сравнения приведены на рисунке 2.

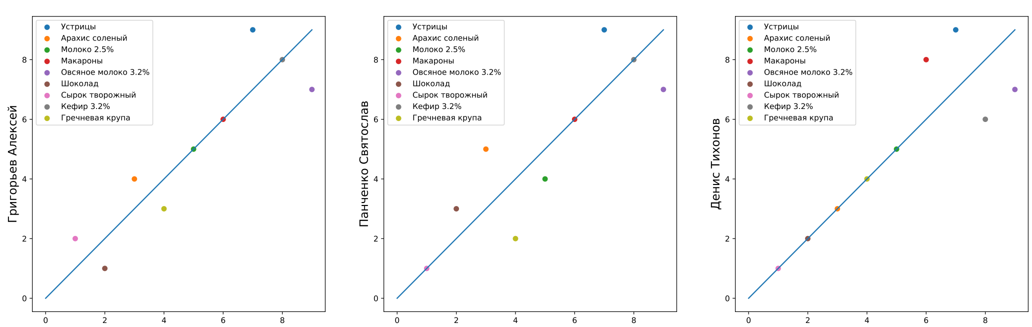


Рис. 2 Сравнение рейтингов

2.3 Предпочтение экспертов

Так же балы рассмотрено решение задачи с предпочтением экспертов. Пусть у нас есть рейтинг экспертов $\mathbf{r} \in \mathbb{R}^K$, состоящий из неповторяющихся оценок от 1 до K. Чем ниже оценка эксперта, тем предпочтительнее его оценки. Чтобы учесть данный рейтиг экспертов, предлагается считать медиану, расширяя матрицы признаков и весов, то есть оценки экспертов использовать как число столбцов для этого эксперта. То есть повторяющиеся столбцы. Например, для эксперта с рейтингом 5, в матрице оценок весов и матрице оценок объектов будет 5 одинаковых столбцов. Взятие медианы по таким матрицам давало бы 'взвешанные средние' $\mathbf{q_0}$ и $\mathbf{w_0}$ с учетом того, что мы работаем в не линейных шкалах. Затем было проведено согласование. Результаты сравнения приведены на рисунке 3.

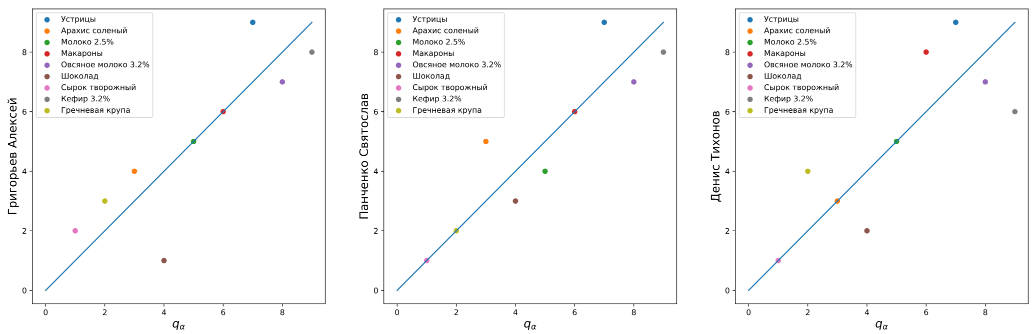


Рис. 3 Сравнение рейтингов

Сравнение рейтингов, полученных без учета и с учетом рейтинга экспертов приведено ниже 4. Так же приведено изменение Q в зависимости от рейтинга эксперта 8. Где Q:

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

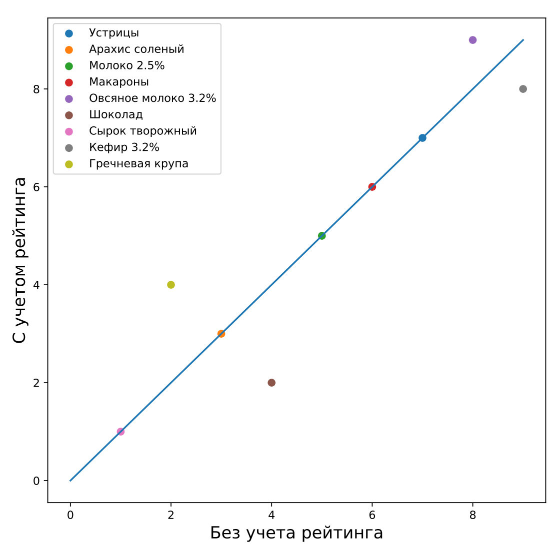
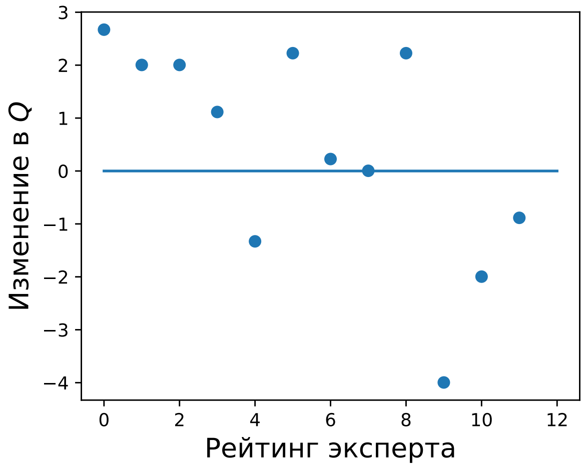


Рис. 4 Сравнение рейтингов



Код: `https://github.com/Intelligent-Systems-Phystech/Ratings/tree/main/Alsahanova%20Nadezhda`