

Вторая задача

Идеи к реализации

- 1) Константа
- 2) Модель

Константа

Для каждой ТД можно сделать оптимальное соотношение количество курьеров на число заказов.

Нужно поварьировать соотношение, чтобы найти такое число курьеров, чтобы опоздания были минимальны.

Модель (1 вариант)

Модель регрессии предсказания delay rate в зависимости от числа курьеров:

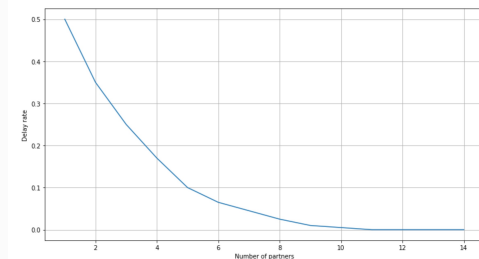
$F(X, \text{orders}, \text{partners}, \text{perc}) = \text{delay rate}$

- ❑ Orders – Количество заказов
- ❑ Partners – Количество курьеров
- ❑ Perc – соотношение курьеров на заказы
- ❑ X – остальные признаки

Для наших предсказаний из первой модели (прогноз числа заказов) давайте поварьируем число курьеров и посмотрим предсказания delay rate.

Посмотрим на кривую и принимаем решение о кол-ве курьеров.

Оптимальное количество курьеров: delay rate < 5%



Модель test

Как лучше построить test выборку:

- 1) У вас есть sample для предсказания:
- 2) Вам нужно его размножить и добавить курьеров и соотношение:

delivery_area_id	data	n_orders
0	1 2021-11-04	11

delivery_area_id	data	n_orders	n_partners	perc
0	1 2021-11-04 14:00:00	11	1	11.000000
1	1 2021-11-04 14:00:00	11	2	5.500000
2	1 2021-11-04 14:00:00	11	3	3.666667
3	1 2021-11-04 14:00:00	11	4	2.750000
4	1 2021-11-04 14:00:00	11	5	2.200000
5	1 2021-11-04 14:00:00	11	6	1.833333
6	1 2021-11-04 14:00:00	11	7	1.571429
7	1 2021-11-04 14:00:00	11	8	1.375000
8	1 2021-11-04 14:00:00	11	9	1.222222
9	1 2021-11-04 14:00:00	11	10	1.100000
10	1 2021-11-04 14:00:00	11	11	1.000000
11	1 2021-11-04 14:00:00	11	12	0.916667
12	1 2021-11-04 14:00:00	11	13	0.846154
13	1 2021-11-04 14:00:00	11	14	0.785714
14	1 2021-11-04 14:00:00	11	15	0.733333

- 3) Теперь, закидываете в модель и получаете delay rate для каждой строки в зависимости от количества курьеров.

Модель (2 вариант)

Модель классификации предсказания $\text{delay rate} < 0.05$ в зависимости от числа курьеров:

$$F(X, \text{orders}, \text{partners}, \text{perc}) = I(\text{delay rate} < 0.05)$$

- ❑ Orders – Количество заказов
- ❑ Partners – Количество курьеров
- ❑ Perc – соотношение курьеров на заказы
- ❑ X – остальные признаки

Для наших предсказаний из первой модели (прогноз числа заказов) давайте поварьируем число курьеров и посмотрим предсказания delay rate (вероятность нулевого delay rate).

Постепенно увеличиваем число курьеров, пока delay rate не станет меньше 5%.

Модель (3 вариант)

Гибрид регрессии и классификации:

$F_1(X, \text{orders}, \text{partners}, \text{perc}) = \text{delay rate}$

$F_2(X, \text{orders}, \text{partners}, \text{perc}) = I(\text{delay rate} < 0.05)$

❑ Ответ: $F_1 * F_2$ **(Как запасной вариант)**

❑ Ответ: $F_1 * (F_2 > \alpha)$ (α подобрать на валидации)