

Классификатор брендингования Яндекс.GO

1. Данное задание включает в себя ряд подзадач, в том числе предлагается изучить возможность trade-off между основными метриками, и одновременно с этим учесть показатель удовлетворённости партнёров.

Основные метрики выглядят следующим образом:

$$Defect\ Rate = \frac{trips\ rated < 4}{all\ rated\ trips}$$

$$Share\ of\ Voice = \frac{Yandex.\ Go\ cars}{all\ branded\ cars}$$

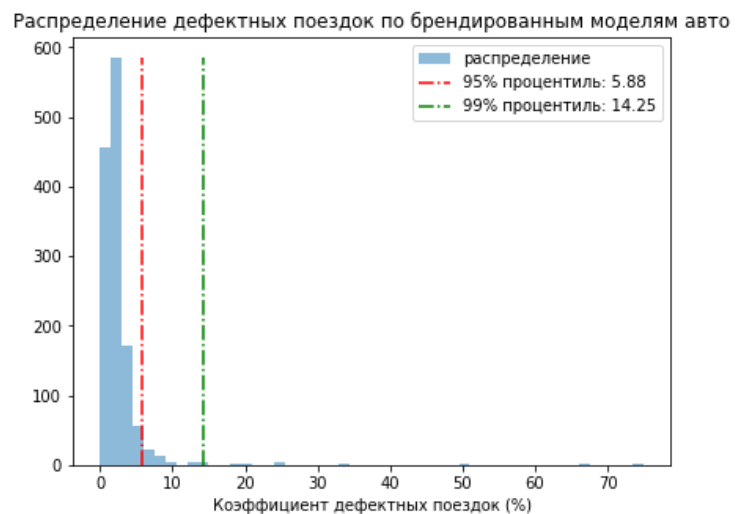
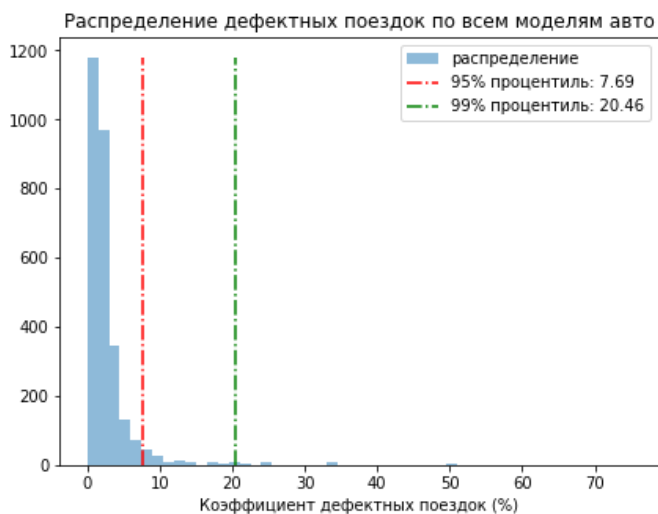
Учитывая специфику наших данных, можно внедрить метрики в дополнение к уже имеющимся, например, такие как:

$$\text{Коэффициент дефектных поездок} = \frac{trips_defect_cnt}{trips_rated_cnt}$$

и

$$\text{Доля "рекламного голоса"} = \frac{car_sticker_cnt}{all_cars_cnt}$$

2. Можно визуализировать распределение дефектных поездок по моделям авто:



Исходя из визуализации можно предположить, что у моделей авто, которые подлежат брендированию сейчас (график справа) defect_rate ниже, однако, возможно, не все модели с низким defect_rate брендируются.

Таким образом, в качестве улучшений можно предложить следующее:

- Уменьшить defect_rate у моделей, подлежащих оклеиванию
- Одновременно учесть показатель share_of_voice
- А также учесть показатель удовлетворённости партнёров

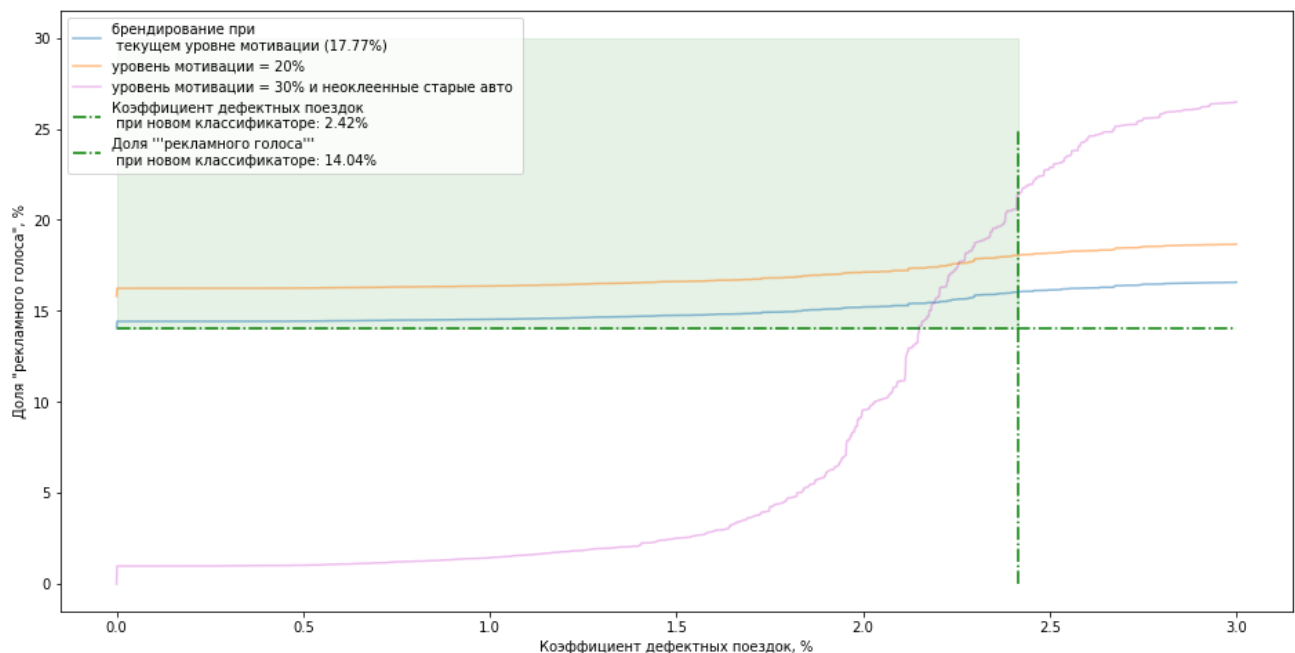
3. Для целей дальнейшего анализа можно внедрить следующую метрику:

$$motivation = \frac{car_sticker_cnt}{car_allowed_to_sticker_cnt}$$

что означает долю тех, кто оклеен, из тех, кому по классификатору можно оклеиваться.

Далее обозначим следующие стратегии оклеивания авто:

- Фиксируем некий defect_rate. Теперь к оклеиванию допускаются те авто, которые допускались по старому классификатору + если у модели авто defect_rate < фиксированного уровня defect_rate. Далее рассчитываем долю таких разрешённых к оклеиванию авто от общего количества авто. Домножаем на motivation при подсчёте.
- То же, что в предыдущем пункте, но незначительно увеличиваем motivation.
- Даже при условии, если новый классификатор делается с нуля (то есть из расчета, что все авто не оклеены), motivation также будет учитываться.



Таким образом, нам необходимо попасть в зелёную часть плоскости для того, чтобы понизить `defect_rate` и повысить показатель `share_of_voice`.

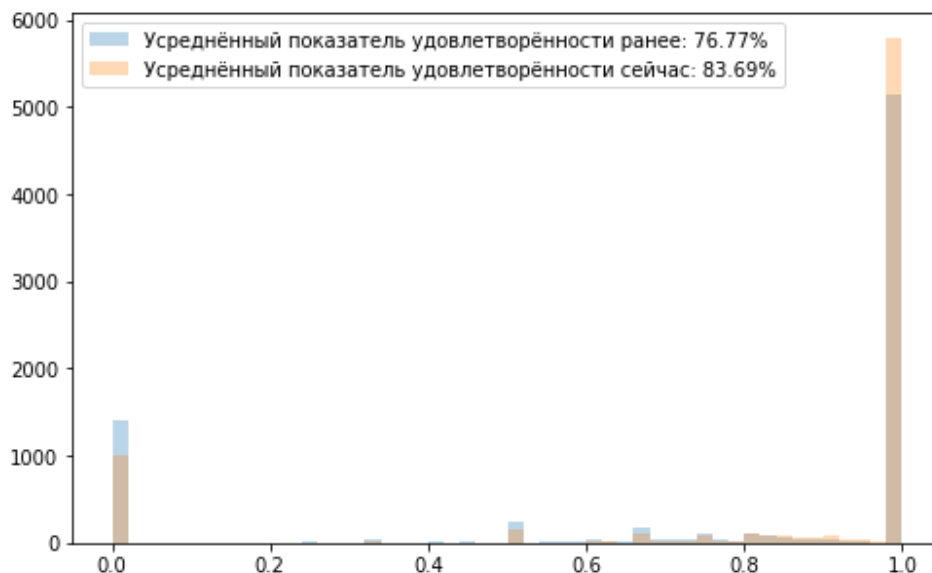
4. Далее можно проанализировать, как такое решение отразится на показателе удовлетворённости партнёров.

Как было сказано ранее, перед нами стоит задача: брендировать те авто, которые допускались к оклеиванию по старому классификатору, + модели авто с `defect_rate < 2%`. Для этого вводим показатель удовлетворённости партнёра:

$$satisfied = \frac{car_allowed_to_sticker_cnt}{all_unstickered_cars_cnt}$$

то есть доля разрешённых к оклеиванию авто из всех не оклеенных.

График распределения этой метрики по партнёрам выглядит следующим образом:



5. Продумывание продуктовых сценариев:

Исходя из результатов анализа можно предложить следующие шаги:

- Модели, разрешённые к брендированию по текущему классификатору, оставляем, но также разрешаем брендирование моделей авто с `defect_rate < 2%`, которые до сих пор не брендировались
- Также важно прорабатывать метрику *motivation* (доля оклеенных из тех, кто подлежит оклеиванию).