

Брендинг машин в такси

1. Постановка целей для нового классификатора

В презентации шла речь о трейд-оффе между следующими метриками:

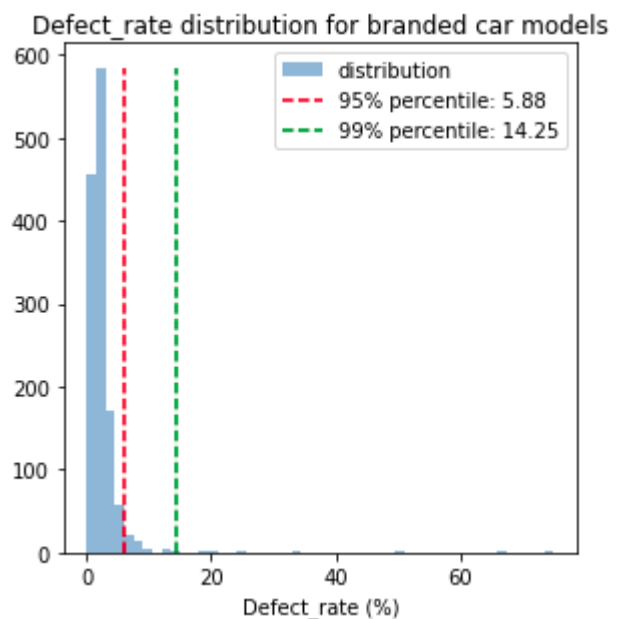
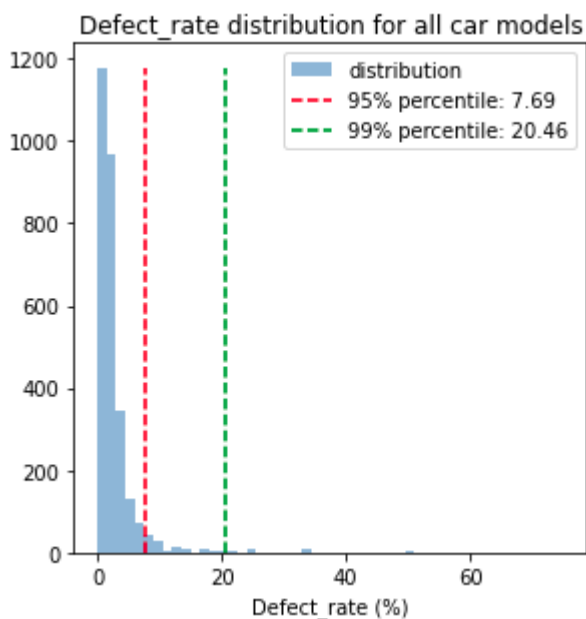
- $ShareOfVoice = \frac{Y_{andexGoCars}}{allBrandedCars}$
- $DefectRate = \frac{tripsRated<4}{allRatedTrips}$

И еще надо как-то учесть партнеров (чтобы они были довольны)

Введем некоторые аналоги для этих метрик, которые можно посчитать по нашим данным:

- $DefectRate = \frac{tripsDefectCnt}{tripsRatedCnt}$ (то есть он в таком виде и считается по данным)
- $ShareOfVoice = \frac{StickerCarsCnt}{AllCarsCnt}$ (доля машин со стикерами от всех машин, которые мы теоретически можем покрасить)

Для начала посмотрим на распределение DefectRate по моделям машин:



Видно, что у моделей, которые брендируются сейчас (график справа), defect rate поменьше. Однако, возможно не все модели с низким defectRateм брендируются (а почему бы их не брендировать).

То есть можно поставить **цели**:

- уменьшить defectRate у разрешенных моделей
- при этом сторговаться с shareOfVoice
- при этом постараться не обидеть никого из партнеров (посмотреть на долю разрешенных к покраске авто из тех, что еще не покрашены -- если она у партнера упала, то грустно)

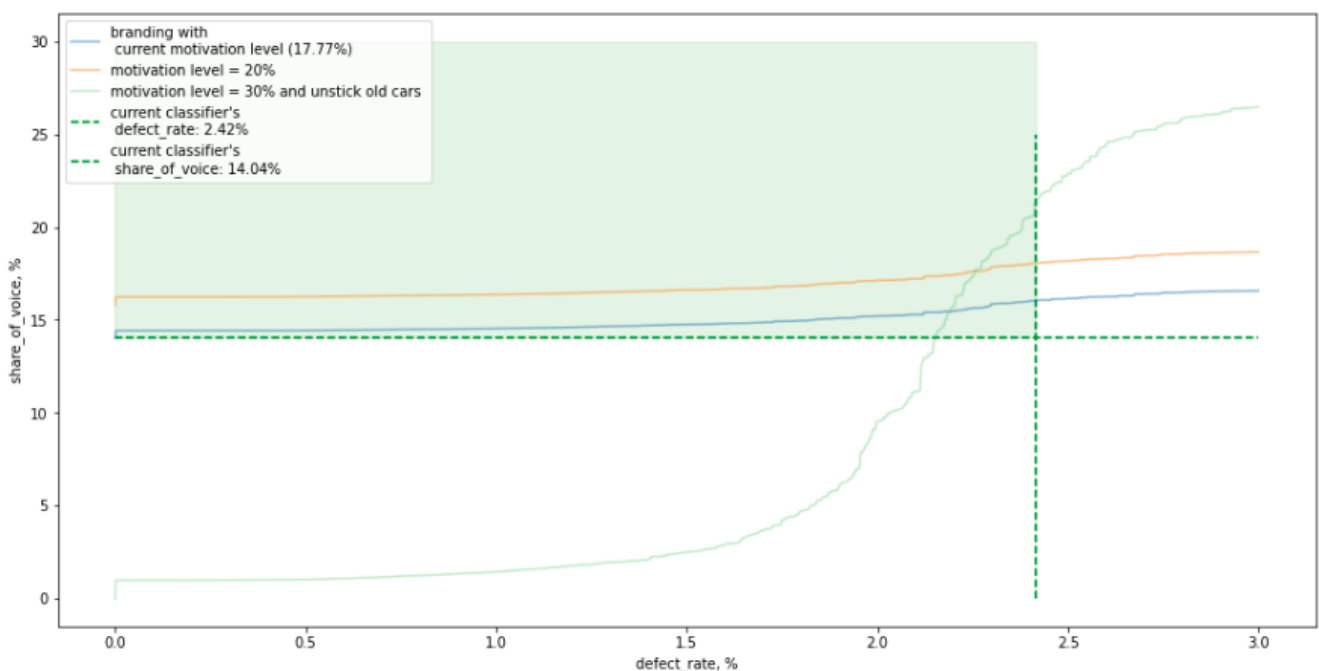
Замечание: конечно, defectRate в том виде, в котором он сейчас, не совсем правильно учитывать при выборе разрешенных моделей машины (больше бы подошло что-то типа $\text{bad_car_model_tag_cnt}/\text{trips_rated_cnt}$). Но там больше нулей, и вроде люди не всегда ставят почему именно им не понравилась машина. В идеале наверно надо было с какими-то весами учитывать и то, и другое

2. Анализ по имеющимся данным

Во-первых, введем такую метрику как

$\text{motivation} = \frac{\text{carStickerCnt}}{\text{carAllowedToStickerCnt}}$, то есть доля тех, кто окрасился, из тех, кому по классификатору можно. Рассмотрим следующие стратегии покраски:

- Фиксируем какой-то defect_rate. Новые машины допускаются, если по старому классификатору допустились бы, или если у их модели defect_rate < фиксированного уровня. Считаем долю таких разрешенных к покраске от всех машин (аналог shareOfVoic-a). Домножаем на motivation при подсчете (пусть он такой же, как был в исходных данных)
- То же, что в предыдущем пункте, но немного увеличиваем motivation
- Представим, что классификатор делается с нуля (то есть сейчас никто не окрашен). Здесь тоже будет учитываться motivation

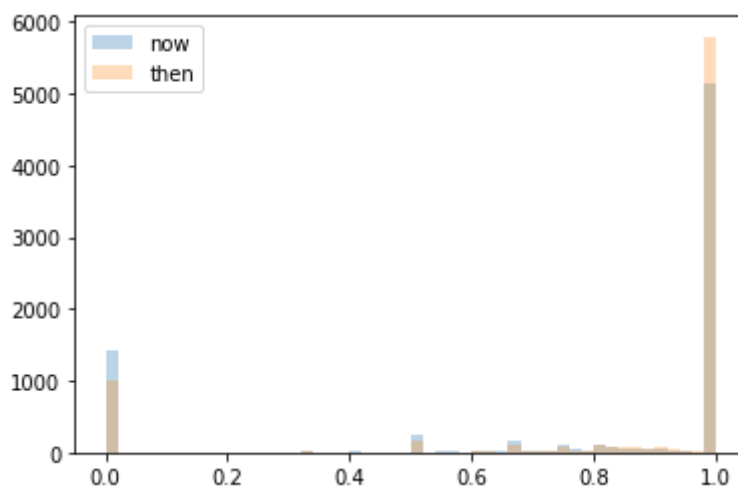


В идеале мы хотим попасть в зеленую четверть плоскости (чтобы понизить defectRate и повысить ShareOfVoice). В целом по графикам кажется, что если мы понизим разрешенный defectRate до 2%, ничего страшного не случится (то есть shareOfVoice поднимется, мы не все новые покрашивания запретили).

Теперь проверим, как такое решение отразится на партнерах. Итак, мы хотим разрешить брендировать то, что было разрешено и еще модели с defectRate-ом < 2%. Введем метрику счастья партнера: $happiness = \frac{allowedToSicker}{alUnstickeredCars}$, то есть доля разрешенных к покраске автомобилей из всех неокрашенных

- mean happiness now: 76.77%
- mean happiness then: 83.69%

График распределения этой метрики по партнерам:



Видимо, было много партнеров с одной машиной, которую раньше было нельзя брендировать, а теперь можно. Ну ладно, хорошо

3. Продумывание продуктовых сценариев

На основе анализа можно предложить такой план действий:

- Модели, которые можно брендировать сейчас -- оставляем, модели с низким defect rate-ом (< 2%), которые еще не брендировались -- разрешаем брендировать
- Видимо, очень важно прокачивать метрику motivation (доля оклеившихся из всех тех, кому это разрешено). Почему важно: тогда share of voice увеличится, а defect rate не станет хуже в рамках наших данных, потому что набор разрешенных моделей не изменился.
- Понятно, что метрика из предыдущего пункта связана с ценой окраски тоже. Данных по стоимости и по скидкам при покраске нет, но из такого прям простого -- можно убедиться, что все водители разрешенных к покраске авто знают, что им можно покраситься (а то вдруг кто-то еще просто не прочитал про субсидии и приоритет при получении заказа).