**Расшифровка по таблице для пунктов 1, 2, 3**

В таблице pictures\_market представлены вердикты для фото (эти фото подойдут для пунктов 1, 2 и 3 из кейса):

* «хорошая инфографика» (одна или несколько понятных и информативных характеристик товара в текстовом или визуальном оформлении, изображение качественное, понятно, что за товар) - "good\_infographics"
* «плохая инфографика» (когда текст без визуального оформления/сложные термины/непонятные символы/только иностранный язык и т.д) - "bad\_infographics"
* «фото без инфографики с хорошим однотонным фоном» - "clean\_photo\_good\_background"
* «фото без инфографики с хорошим интерьерным фоном»» - "clean\_photo\_image\_background"
* «фото без инфографики с плохим фоном» - "clean\_photo\_bad\_background"
* «не относим ни к плохой/ни к хорошей инфографике» - "other\_infographics"
* «не относим к фото товара без инфографики» - "clean\_photo\_other\_background"

Данная разметка на вердикты сделана внешними толокерами (здесь можно посмотреть инструкцию по разметку Инфографики - <file:///Users/eastar/Downloads/Инфографика.html>, а здесь можно посмотреть инструкцию по разметке Фона у фото без инфографики - <file:///Users/eastar/Downloads/Фон.html>) и иногда похожие/аналогичные картинки могут лежать в разных вердиктах. Это связано с тем, что в целом бывает сложно провести формальную грань между некоторыми вердиктами.

Чаще всего можно заметить, что картинки «плавают» между:

* «хорошей» и «плохой» инфографикой (большую роль здесь играет визуальное восприятие при разметке/насколько «продающе» выглядит фото / не слишком ли ляписто/ не слишком ли мелкий шрифт)
* «фото без инфографики с хорошим интерьерным фоном» и «фото без инфографики с плохим фоном» (в данном случае тонкая грань проводится тоже на уровне «красивости» интерьерного фона – когда это профессиональная съемка на интерьерном фоне – считаем хорошим фоном / а когда фото выглядит «грязным», неаккуратным, непрофессионально отснятым – это уже плохой фон)

Также в разметке приложены фото с вердиктами «не относим ни к плохой/ни к хорошей инфографике» и «не относим к фото товара без инфографики».

Как правило,сюда попадали случаи, когда:

* на фото товара есть вотермарка (это водяные знаки с названием магазина, пример - <https://www.visualwatermark.com/images/pdf/watermark-example.webp>)
* на фото товара указаны условия продажи этого товара конкретным магазином (например: «Гарантия 1 год», «Доставка 2 дня», «Лучшая цена», «Акция»)
* на фото вообще нет товара, там только текст
* толокеры не смогли определить разницу между «хорошей» и «плохой» инфографикой или «хорошим» и «плохим» фоном – и для подстраховки отправляли такие фото в вердикт «Другое» (чтобы не портить качество ответов в основных вердиктах на разбивку хороших/плохих фото)

Вы можете использовать эти вердикты про «другое» на свое усмотрение в рамках дообучения модели.

При принятии задания по классификации товаров на инфографику/не инфографику может попасться пример – когда отдадим на входе картинку вообще без фото товара – и ожидается, что в данном случае модель не отдаст такому фото ни «хорошего», ни «плохого» вердикта.

**Расшифровка по таблицам для пункта 4**

В папке «Параметры по фото» есть 10 таблиц – 1 таблица на 1 товарную категорию.

В каждой категории операторы Маркета извлекали по фото на входе определенные характеристики товара (названия характеристик находятся в первой строке, а в остальных строках заполнены значения этих характеристик; почитать поподробнее про сами характеристики и их значения можно из доп. таблички «Характеристики в категориях»; искать параметры по полю category\_id + xsl\_name).

Если оператор по фото на входе не смог заполнить значения в параметре – оставалась пустота в ячейке. В файлах можно найти строки, где по фото не удалось заполнить вообще ни одного параметра.