

Описание

Вам прислали кропы знаков для классификации. Однако, вы заметили, что достаточно много знаков на изображениях сильно обрезаны и портят вам тренировку. Поэтому вы решили обучить еще одну сеть фильтровать такие изображения. Но датасета для тренировки у вас еще нет. Поэтому встала задача собрать такой датасет.

Цель

Получить разметку, в которой указано целиком ли знак на изображении или нет.

Требования к проекту

Должны быть следующие варианты ответа, что знак:

- **Полностью на изображении** с лейблом **ok**
- **Обрезан** с лейблом **bad**
- **Нет знака** с лейблом **empty**
- **Ошибка загрузки** с лейблом **error**

В самом проекте должно быть три пула:

1. Обучающий
2. Экзаменационный
3. Основной

Требования к обучающему пулу

- В названии указываем, что это обучение
- В блоке *Настройки прохождения* указать *Полное прохождения*

Требования к экзаменационному пулу

- **Перекрытие** - любое (можно оставить 3)
- **Цена** - 0
- В экзаменационный пул должны попадать только те исполнители, которые прошли обучение (порог можно указать 70%)
- В названии должно быть указано, что это экзамен
- **Тип пула** - Экзамен (в блоке *Дополнительные настройки*)

- Делаем фильтр на русскоязычных исполнителей
- Создаем навык **Целиком ли знак на изображении**
- В блоке *Контроль качества* в *Контрольные задания* присваиваем в наш навык процент правильных ответов (нужно учитывать все задания из экзамена)
- При загрузке tsv файла в умном смешивании на все основные и обучающие задания ставим 0. На контрольные задания ставим число контрольных заданий

Требования к основному пулу

- **Перекрытие** - 3
- **Цена** - 0.02 за страницу
- В блоке *Аудитория* фильтруем по навыку *Целиком ли знак на изображении* - порог 80
- Оставляем только русскоязычных исполнителей
- По контрольным заданиям блокируем тех, у кого процент правильных ответов меньше 90 (учитываем последние 10)
- Остальные параметры настраиваем по своему усмотрению