

#### Обогащение данных:

- увеличить размерное пространство признаков
- собрать данные о всевозможных потенциально значимых объектах в какой-то близости от целевой точки (метро, школы, больницы и тп)
- разработать новые признаки из существующего адреса (лучший процессинг, чем в бейзлайне, добавить доп признаки, такие как улица, дом и тп)

#### Потенциальные модели:

- Catboost (возможны другие имплементации бустингов, покажет сравнительный анализ)
- Random forest (с предобработкой категориальных фичей)

#### Архитектура приложения:

Все приложение состоит из 3-х основных узлов:

- 1) ML модель
- 2) Парсер данных
- 3) Интерфейс ввода-вывода

В роли интерфейса ввода-вывода будет выступать Телеграм бот, написанный на python и выполняемый запуском соответствующего .py скрипта.

Парсинг данных и обсчет модели будет осуществляться один раз в неделю, либо при каждом развертывании и запуске приложения.

Все скрипты, сопутствующие файлы и прочее будет упаковано в docker образ и в дальнейшем развернуто удаленно.