

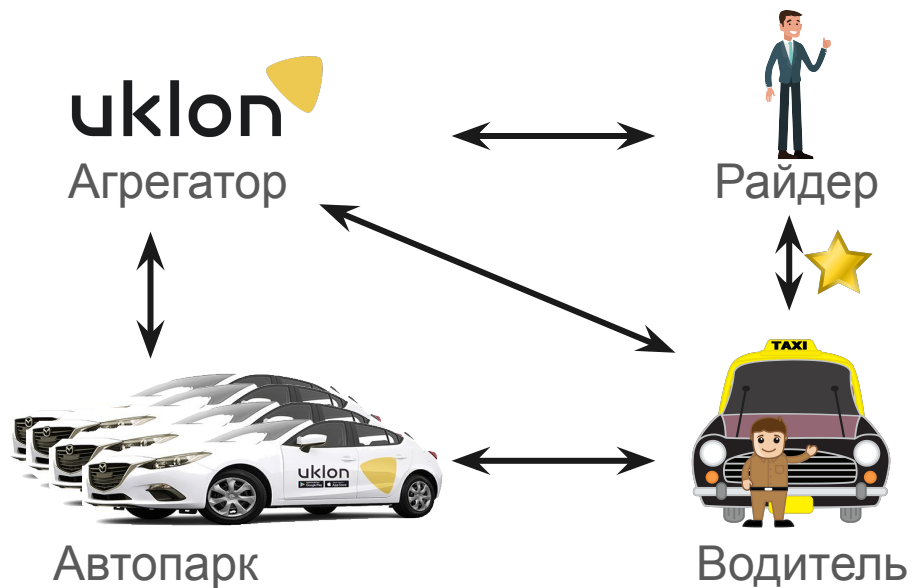


Graph optimization

“Distrust Level”:

Геллер Михаил, Морозов Константин, Мхитарян Нарек, Погребная Александра

Постановка задачи



Поощряем хороших
водителей



Идеи



Процент от выполненного заказа



Выбор выходного дня



Карта прогноза спроса



Заказы по пути домой



Данные

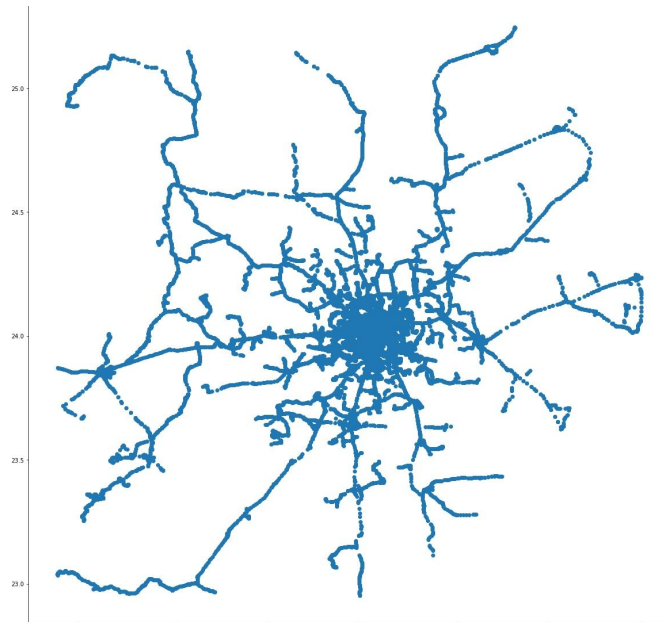
- Вычищаем: выбросы, тренд, сезонность
- Считаем рейтинг водителей
- Добавляем еще признаки



Адреса водителей



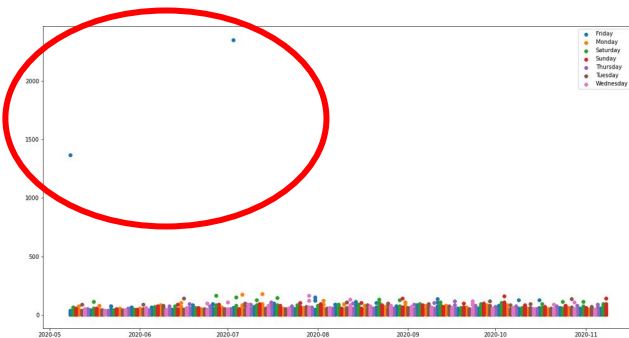
История заказов



Карта города

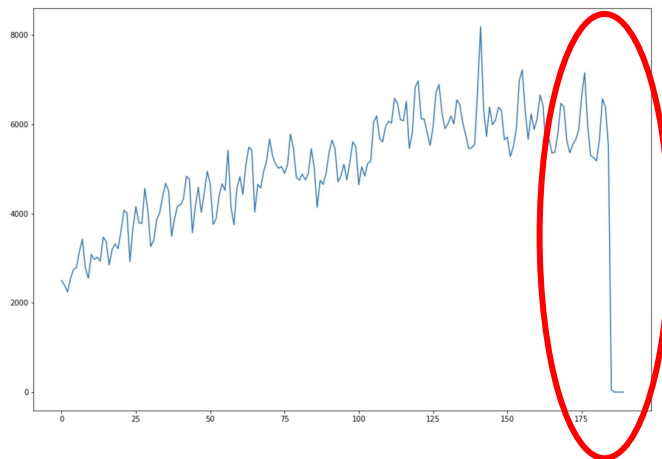
Данные. Очистка

Выбросы

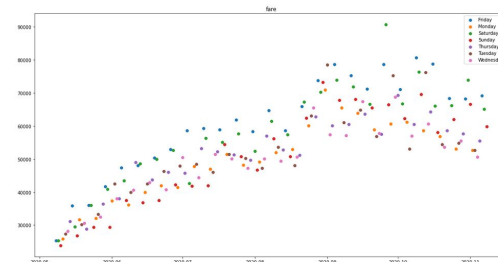


	driverID	dropoffTime	dropoff_lat	dropoff_lon	fare	pickupTime	pickup_lat	pickup_lon	rideRating	pickup	dropoff	week_day
326112	732302	7/3/2020				7/3/2020						
		15:54	49.843609	24.031656	2349.71035	15:21	49.81884	23.92687	NaN	2020-07-03 15:21:00	2020-07-03 15:54:00	Friday
734967	3072419	5/8/2020				5/8/2020						
		22:12	49.812890	24.065860	1367.47230	21:31	49.85927	24.01834	NaN	2020-05-08 21:31:00	2020-05-08 22:12:00	Friday

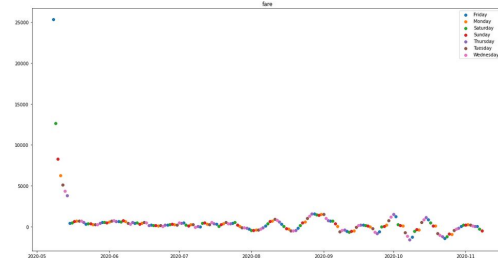
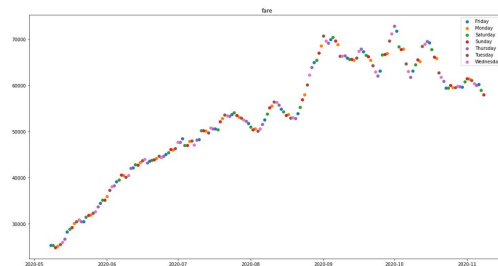
Нехватка данных



Недельная сезонность



Тренд





Данные. Рейтинг



Оценки

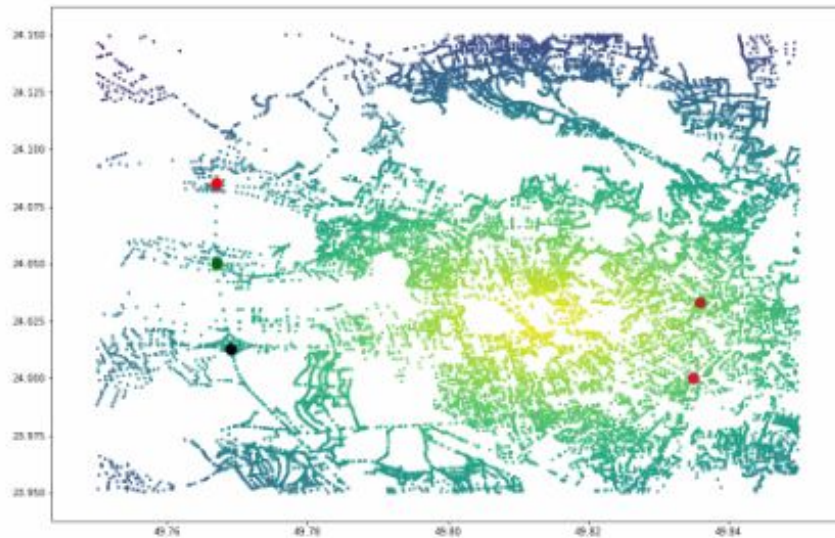


Опыт

Данные. Карта города

Учитываем:

- расстояние от дома водителя до центра города
- среднюю скорость его поездок



$n \log(n)$ на Дейкстру

Машинное обучение. Идея

baseline

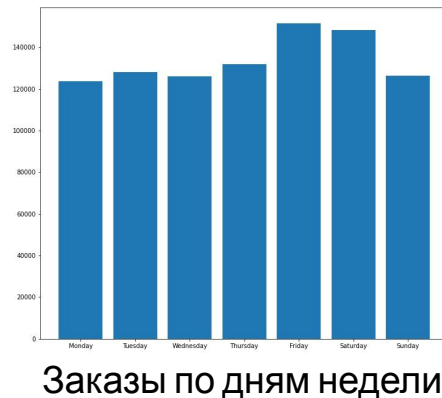
выходные - дни
с наименьшим
числом заказов

driver	week day 1	week day 2
1	?	?
2	?	?
3	?	?

Алгоритм распределения

1. баланс по дням
2. выходные дни
предлагаются с учетом
рейтинга водителя
3. учет водителей, живущих
далеко от центра города

fit
predict → \$ Доход
водителя



Машинное обучение



drivers'
features



Yandex
CatBoost



20* минут

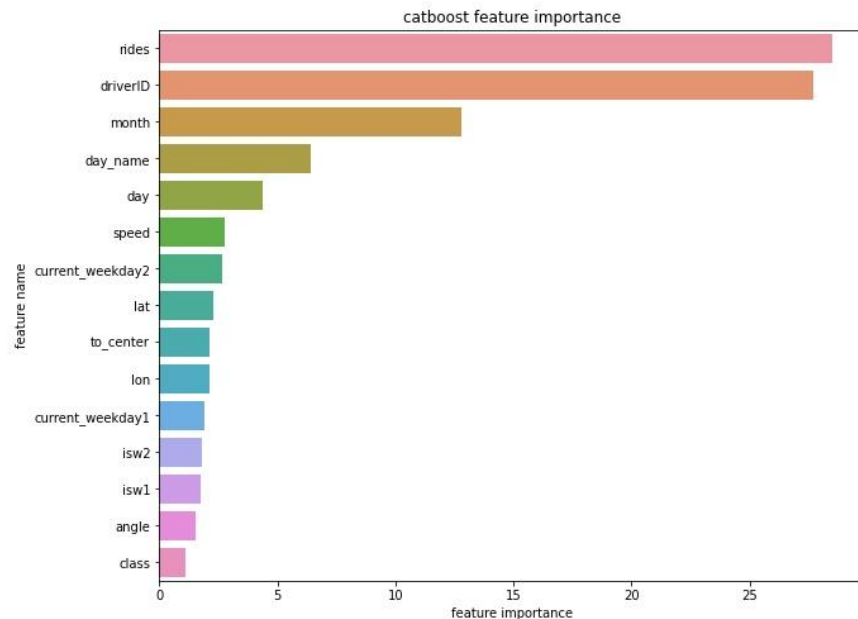


*на 5000 эпох, 200000 строк, 16 столбцов

Результаты

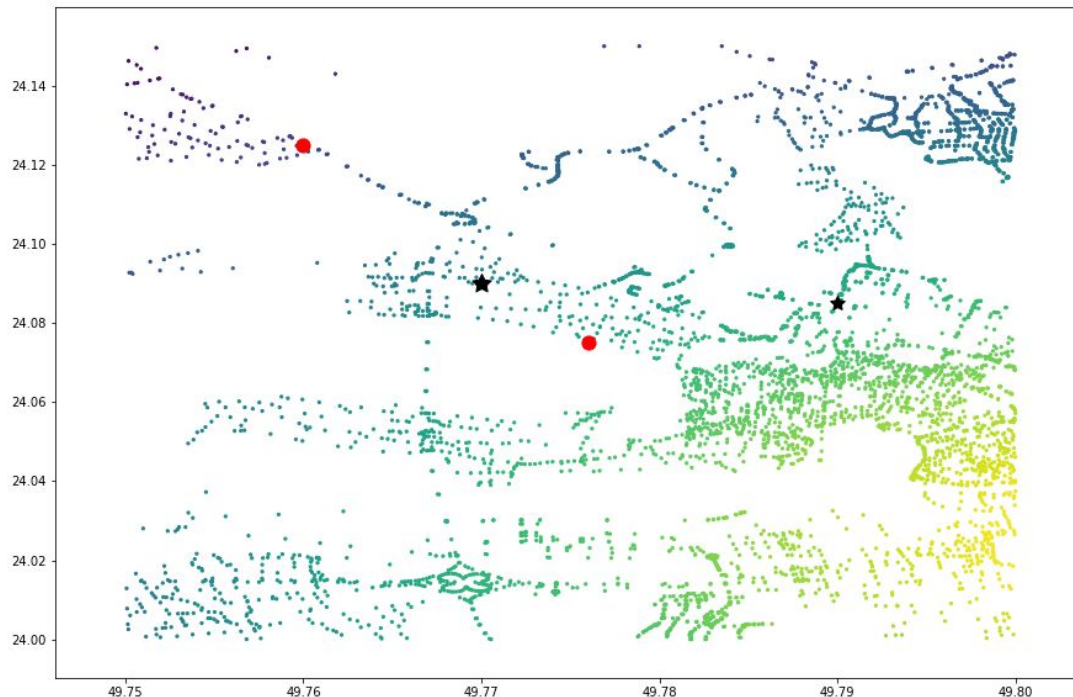
Суммарный доход первоклассных водителей вырос на **16%**.

Суммарная прибыль **не уменьшилась**.



Идея получения реальной прибыли

Будем учитывать
затраты на
энергоресурс!




Спасибо за внимание!

