



STROTСКА

× ×

× ×

× ×

× ×

**РАЗРАБОТКА ПЕРСОНАЛЬНОЙ СТАТИСТИКИ
ОРГАНИЗАЦИИ ПОСТАВЩИКА**

× ×

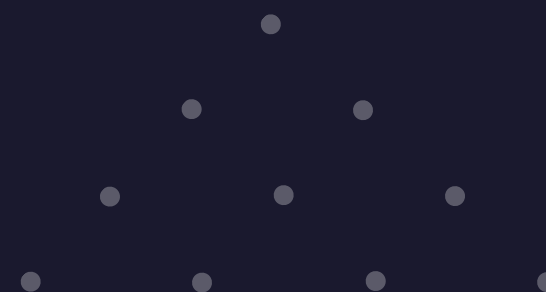
× ×

× ×

× ×



TenderHack



Этапы работы



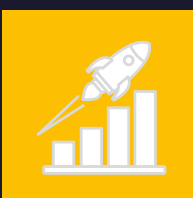
Предобработка данных

Поиск аномалий в данных, удаление выбросов, объединение таблиц, извлечение дополнительных признаков. Feature engineering: выделение категорий товаров методами TF-IDF и word2vec.



Настройка Superset

Организация superset-сервера с помощью Docker, проверка функционала, создание БД, создание графиков с помощью SQL Lab.



Аналитика

Выделение ключевых метрик, создание дашбордов с учётом предиктивной и проактивной аналитики.

24K+

УДАЛЕНИЕ АНОМАЛЬНЫХ СТРОК

- Цена закупки 0 руб, но в контракте >0
- Дублирующиеся id контракта
- Даты контрактов 1970.01.01

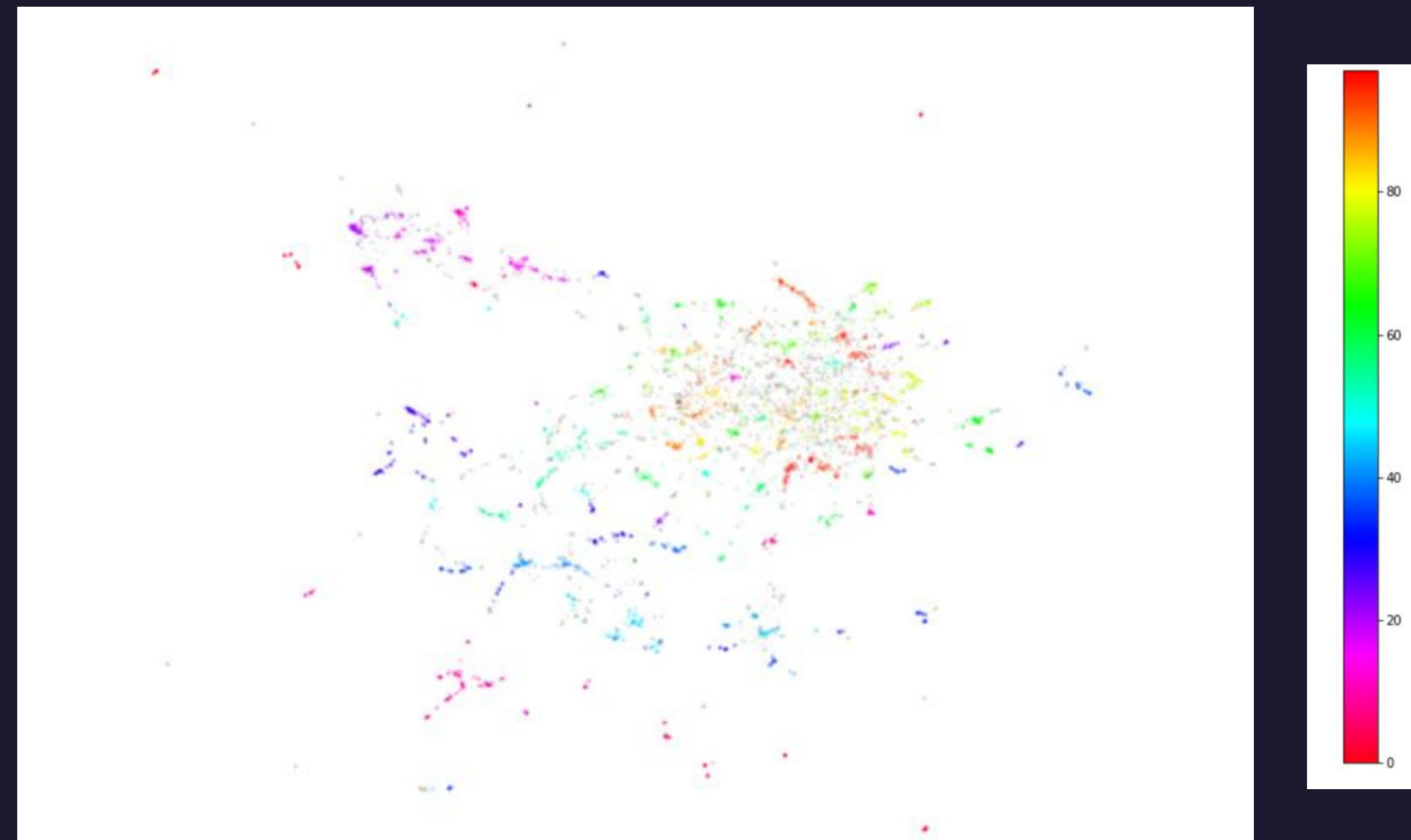
4

FEATURE ENGINEERING

- Категория товара/услуги (поиск близости)
- Время от публикации заявки на закупку до заключения контракта
- Падение стоимости заказа (в процентах)
- Регион доставки (форматирование поля)
- Ценовые категории



ПРЕДОБРАБОТКА ДАННЫХ





Extracting CATEGORICAL FEATURES

Применение методов NLP для
определения категории товара.



ШАГ 01

Извлечение категорий из
справочника.



ШАГ 02

Кластеризация.



ШАГ 03

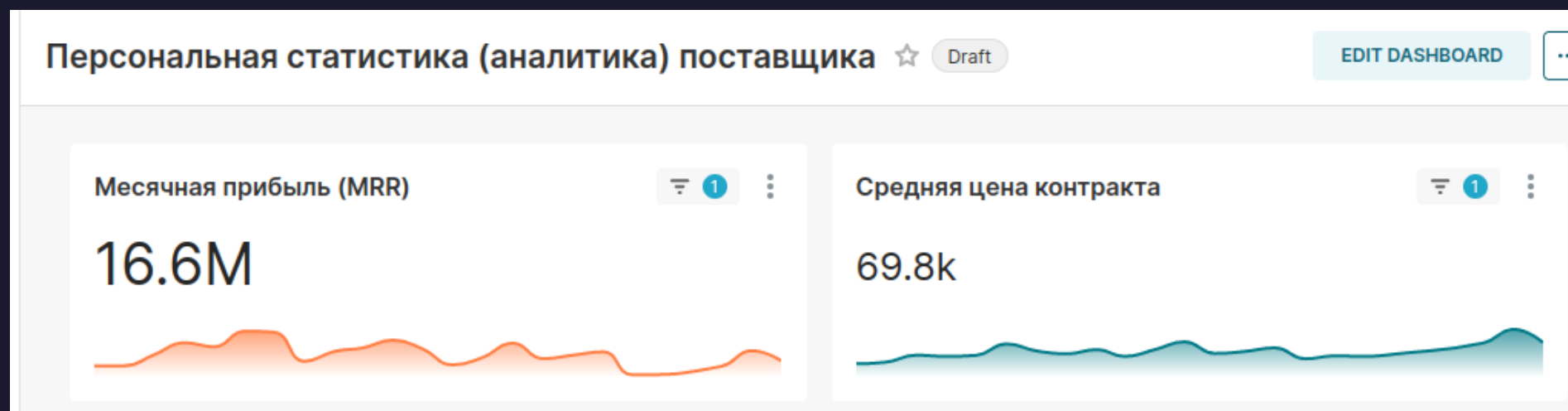
Применение word2vec.

Диагностическая аналитика



Сегменты

Распределение, динамика распределения



Деньги

Прибыль, динамика прибыли, средняя цена контракта, линия тренда



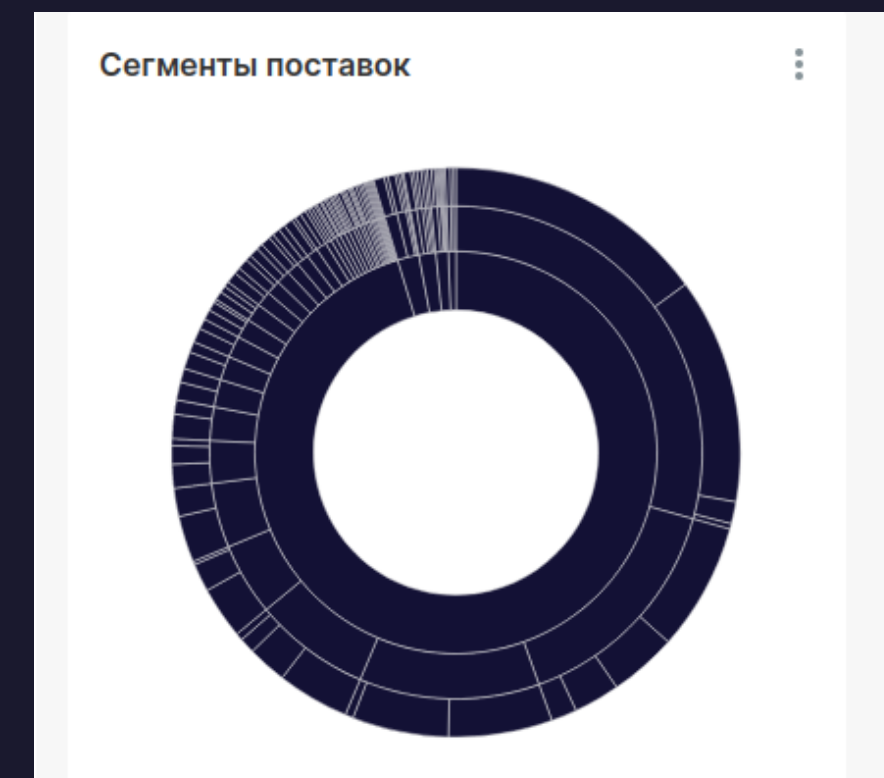
Регионы

Распределение



Люди

Конверсия



Проактивная аналитика

01

Рекомендация по расширению рынка сбыта в свободные от конкуренции регионы

Регионы с минимальной конкуренцией

Search

49 records...

Регион	Коэффициент конкуренции
Свердловская область	0
Республика Ингушетии	0
Ярославская область	0
Красноярский край	0
Волгоградская область	0
Вологодская область	0
Костромская область	0
Республика Татарстан	0
Белгородская область	0

02

Рекомендация по расширению сферы деятельности

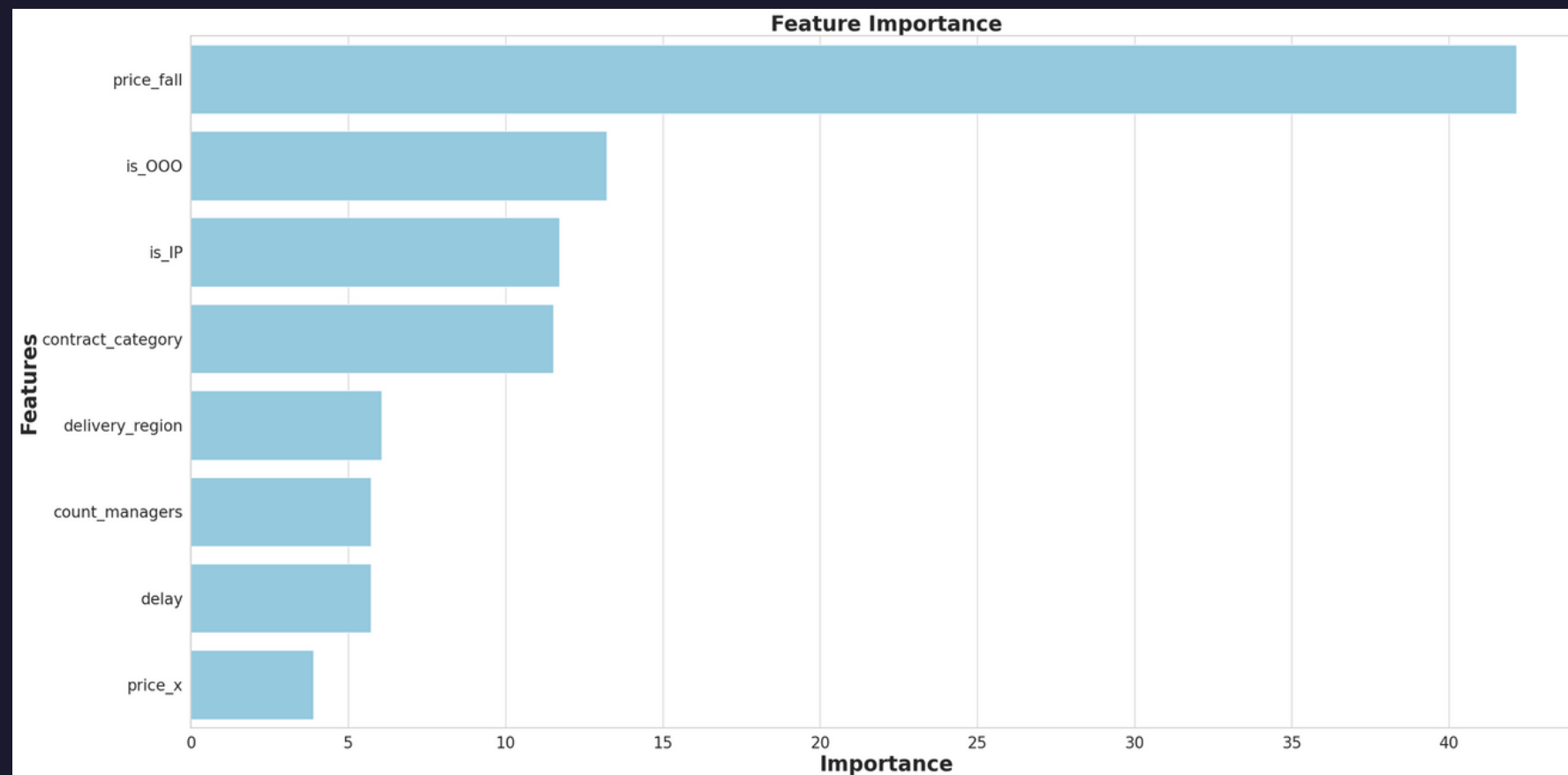
Категории с минимальной конкуренцией

Show 20 entries

theme	cr
Оптовая торговля продуктами питания и напитками	0.54
Услуги по производству и доставке химической продукции	0.72
Бытовые товары	0.96
Техническая розничная торговля	1.01
Бумажная промышленность	1.06
Торговый договор	1.08
Оборудование и электротехнические услуги	1.08
Услуги связи и информационные услуги	1.22
Услуги по транспортировке нефти, газа и воды и снабжению	1.24
Финансовые и страховые услуги	1.25

Предиктивная аналитика

01 CatBoostClassifier → Feature Importance



02 Прогноз динамики дохода





СПАСИБО ЗА ВНИМАНИЕ

Идём смотреть дашборд)

