

Хакатон Т1

Оптимизация сети из 250 клиник

Жизнякова Полина ПМ25-2



Подход и методология

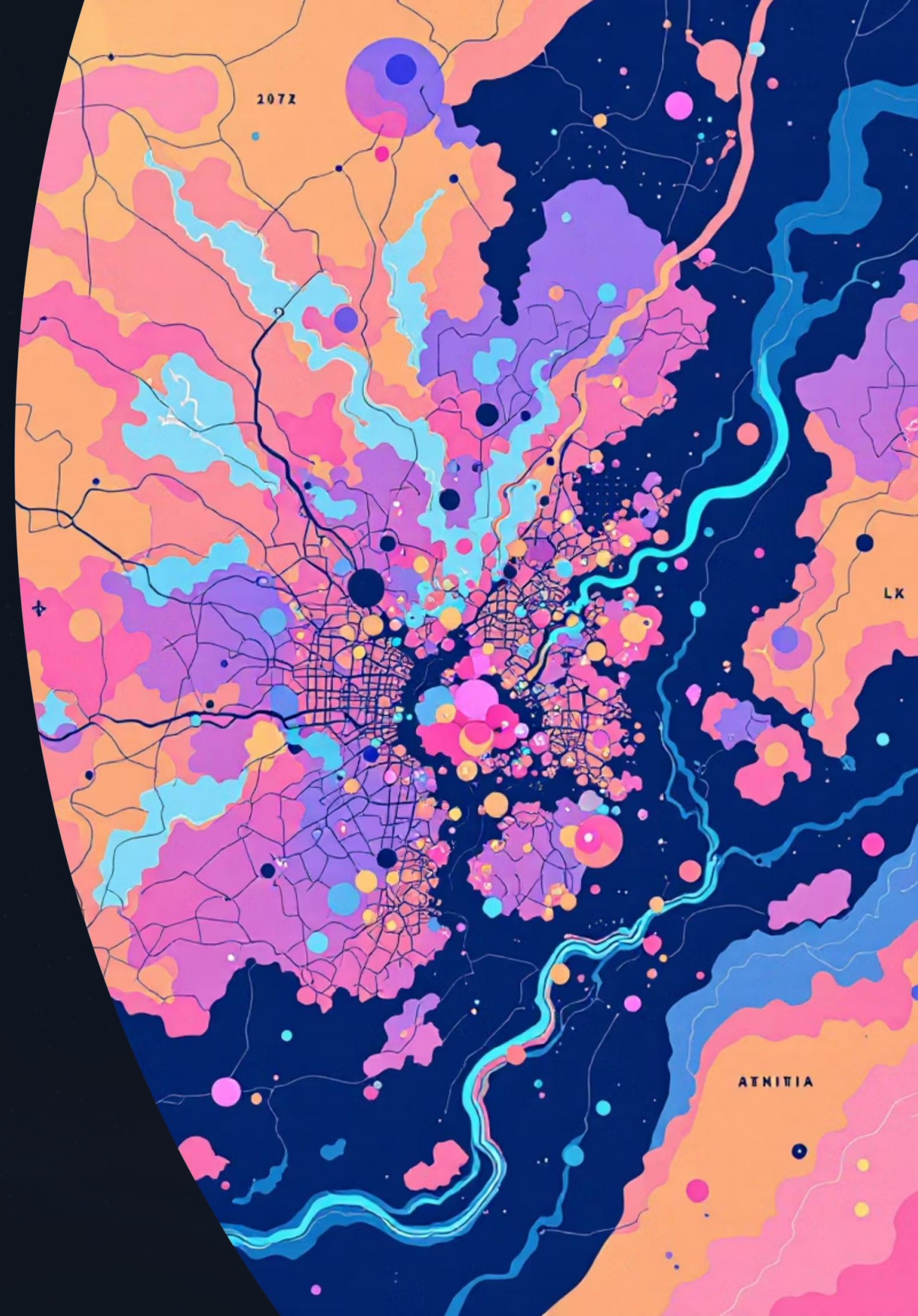
3-этапный алгоритм решения:

- 1. Кластеризация данных (географическая группировка)**
 - Автоматическое разделение на 250 кластеров
 - Равномерное покрытие территории
- 2. Борьба с перегрузкой**
 - Проверка лимита 1000 пациентов
 - Перераспределение клиентов между клиниками
- 3. Оптимизация с учетом фич**
 - Учет 6 социальных характеристик районов
 - Локальная оптимизация положения клиник



Почему именно DBSCAN?

- Автоматически определяет количество кластеров
- Учитывает плотность населения
- Избегает выбросов



Учет социальных факторов

6 ключевых фич из данных:

- **Плотность населения** — где больше всего пациентов
- **Доступность транспорта** — приоритет метро и автобусов
- **Возрастная структура** — внимание к пожилым пациентам
- **Близость к паркам** — экология и здоровье
- **Уязвимые группы** — социальная справедливость
- **Социальная инфраструктура** — общая развитость района

Как используем:

- **Анализируем средние значения фич для каждой клиники**
- **Преобразуем в единую метрику качества (LCS)**



Локальная оптимизация: как выбираем позицию

Для каждой клиники перебираем 25 кандидатных положений в радиусе 1 единицы и вычисляем:

- ΔTTT — изменение среднего времени в пути
- LCS — качество расположения (чем выше, тем лучше)

$Score = \Delta TTT + (100 / (LCS + 1)) * 0.1$ — выбираем позицию с минимальным Score.



Техническая реализация и производительность

Библиотеки:

- **math** — вычисление расстояний
- **csv** — обработка данных

Ключевые функции:

assign_clients_to_clinics() — распределение пациентов

fix_overloaded_clinics() — исправление перегрузки

optimize_with_features() — оптимизация с фичами

calculate_metrics() — вычисление всех метрик





Что делает решение уникальным?

- **Автоматическая кластеризация — не нужно задавать количество кластеров**
- **Гарантия отсутствия перегрузки — строгое соблюдение лимита 1000**
- **Баланс двух критериев — время в пути + качество расположения**
- **Интерпретируемость — прозрачные метрики и логика**
- **Практичность — готовое решение за 2 минуты работы**