



**ДЕПАРТАМЕНТ ЖИЛИЩНО-  
КОМУНАЛЬНОГО  
ХОЗЯЙСТВА  
ГОРОДА МОСКВЫ**

## **КОМАНДА «СЕМЕЙНАЯ»**

**Задача №2. «Рекомендательный сервис прогнозирования  
возникновения технологических ситуаций»**

# КОМАНДА «СЕМЕЙНАЯ»



ПРОЕКТ  
МЭРА  
МОСКВЫ



ДЕПАРТАМЕНТ  
ПРЕДПРИНИМАТЕЛЬСТВА  
И ИННОВАЦИОННОГО РАЗВИТИЯ  
ГОРОДА МОСКВЫ



РАЗВИТИЕ  
ЧЕЛОВЕЧЕСКОГО  
КАПИТАЛА



ЛИДЕРЫ  
ЦИФРОВОЙ  
ТРАНСФОРМАЦИИ

## О команде

- Г.Москва
- 2 участника
- Капитан команды – Андрей Глинский

## Наименование задачи:

Задача №2. «Рекомендательный сервис прогнозирования возникновения технологических ситуаций».

## Описание решения:

Анализ паттернов водопользования (час, день недели: погода, выходные, тип потребителя, сезон и т.д.) и выявление трендов и аномалий для всех узлов графа

Анализ аномалий в показателях счетчиков в узлах графа

Использование CNN/Transformers+полносвязный регрессор (1-й этап) и GCN (графовые сверточные сети) на 2 этапе



## Как вы планируете дальше использовать или развивать ваше решение:

Первый этап - построение графовой модели системы, подготовка исходных данных для обучения GCN/GNN, создание прототипа расширенного сервиса для сегмента сеит Мосводоканала (пилот)

Второй этап – работающие базовая модель и сервис.

Третий этап – Использование LLM для обработки отчетов, журналов аварий и ремонтов, обращений граждан, создание интегрированной системы поддержки диспетчеров (агент-помощник, регламенты, помощь в оформлении отчетов и переписки)

Четвертый этап -добавление данных от других приборов Мосводоканала, интеграция данных других департаментов (строительства, энергетики и т.д.),

# Команда «СЕМЕЙНАЯ»



ПРОЕКТ  
МЭРА  
МОСКВЫ



ДЕПАРТАМЕНТ  
ПРЕДПРИНИМАТЕЛЬСТВА  
И ИННОВАЦИОННОГО РАЗВИТИЯ  
ГОРОДА МОСКВЫ



РАЗВИТИЕ  
ЧЕЛОВЕЧЕСКОГО  
КАПИТАЛА



ЛИДЕРЫ  
ЦИФРОВОЙ  
ТРАНСФОРМАЦИИ

## Краткая история команды:

Команде ровно 40 лет, есть 2 дочерние (сыновние) команды.

Участники являются многократными победителями и призерами региональных и международных хакатонов (Цифровой прорыв, Атомхак, Нефтекод и т.д.)

В ЛЦТ команда «Семейная» - 1 место в Аэротон-2024  
Опыт: 25 лет в нефтегазе и независимой инспекции,  
30 лет - в научной и преподавательской работе в МГУ

## Почему вы выбрали именно эту задачу из предложенных на хакатоне?

Это масштабная задача в консервативной области, решение имеет большой потенциальный эффект для города и, в перспективе, для страны, а также в совершенно несвязанных рынках/отраслях.

Мы - команда со знаниями в реальном производстве, опытом продвижения новых технологий и IT решений и с пониманием особенностей работы больших структур и опытом в управлении изменениями в организации

## С какими основными сложностями или вызовами вы столкнулись и как их преодолели?

Особенности большой структуры, где один департамент/подструктура не сильно заинтересована в том, что делает руководство или другая подструктура – с этим надо просто жить)). Нет данных!

Но, зато есть возможность предложить более широкий подход!

# КОМАНДА «СЕМЕЙНАЯ»



**Андрей  
Глинский**

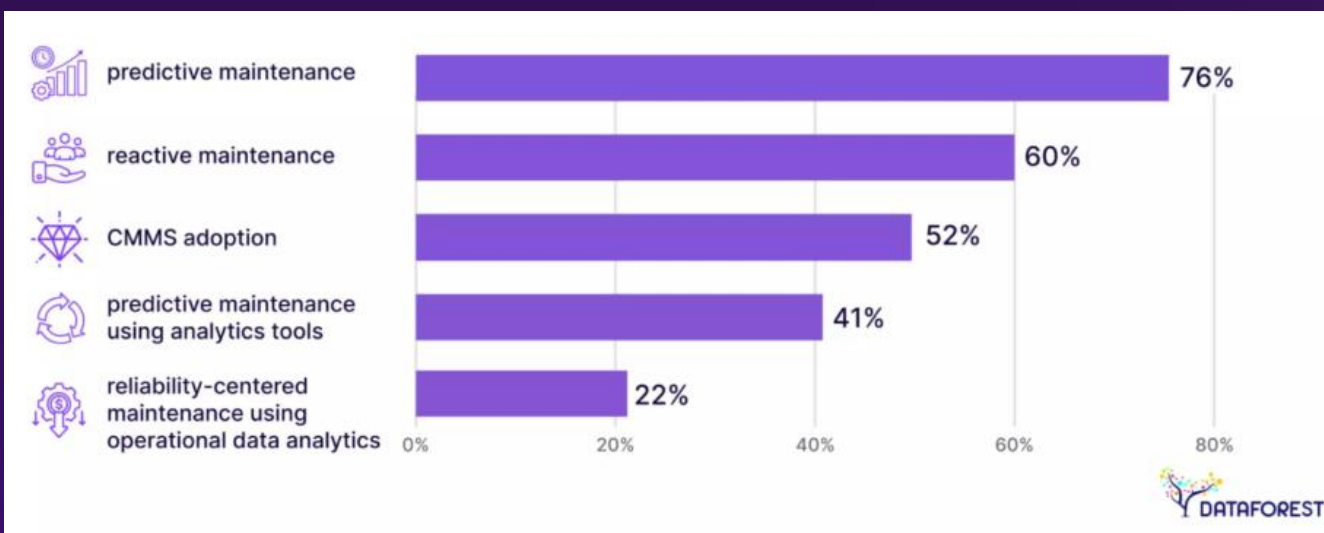
- ML, Business Analytics, Python
- @AI\_glinsky



**Нелли  
Глинская**

- Моральная поддержка
- UI/UX

**Переход от реактивной к проактивной модели – 15-20% снижения потерь за каждый этап внедрения**  
**Реально достижимый экономический эффект от орг. мер / цифровизации– 800-1000 млн. руб. в год**



## • Потери денег

- Суммарный годовой ущерб от отказов трубопроводов Московского водопровода составил в 2011 году более 3873 млн.р.

## • Потери воды

- Не менее 10-15% потерь
- Экологические риски – промоины, затопления, повышенная нагрузка на источники воды и канализацию
- Риски порчи имущества и оборудования

## • Резервы улучшения городской среды

- Качественнее вода
- Удовлетворенные жители и бизнес
- Москва – лучший город Земли!**

# КОМАНДА «СЕМЕЙНАЯ» – ВИДЕНИЕ ПРОБЛЕМЫ



Проблем – 4 одновременно.

- 1) Реактивный подход к обслуживанию и аварийности – неэффективный
- 2) Завышенные потери воды (>10-15% потерь)\*
- 3) Есть значительные резервы в энергоэффективности, но нужны цифры и факты (частотные привода, оптимизация потоков)
- 4) Человеческий фактор – нагрузка на диспетчеров и квалификация специалистов

\* Схема водоснабжения и водоотведения г. Москвы до 2025 года.

Текущие подходы в мире:

Статистические и вероятностные модели, моделирование гидравлики, машинное обучение в области аналитики аварийности, предсказательные модели.

GNN - появляются

Нет идеальной модели в связи со сложностью задачи

Крупные операторы – консервативные структуры

Решение:

- 1) Глубокий анализ потребления воды в зависимости от времени/ места/ погоды/ типа потребителей, выявление ключевых параметров (временные ряды) и выявление аномалий на лету
- 2) Графовые нейронные сети, моделирующие всю систему Мосводоканала для предсказания инцидентов и оптимизации сети
- 3) Чат-бот ассистенты и большие языковые модели в помощь персоналу (регламенты, техдокументация и работа с отзывами граждан и отчетностью)



# СТЕК ПРИМЕНИМЫХ ТЕХНОЛОГИЙ



ПРОЕКТ  
МЭРА  
МОСКВЫ



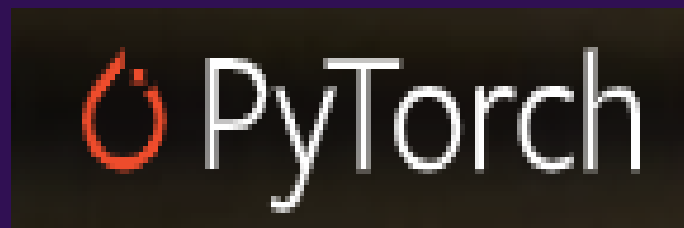
ДЕПАРТАМЕНТ  
ПРЕДПРИНИМАТЕЛЬСТВА  
И ИННОВАЦИОННОГО РАЗВИТИЯ  
ГОРОДА МОСКВЫ



РАЗВИТИЕ  
ЧЕЛОВЕЧЕСКОГО  
КАПИТАЛА



ЛИДЕРЫ  
ЦИФРОВОЙ  
ТРАНСФОРМАЦИИ



# РЕШЕНИЕ – ИДЕОЛОГИЯ И ПОДХОДЫ



ПРОЕКТ  
МЭРА  
МОСКВЫ



ДЕПАРТАМЕНТ  
ПРЕДПРИНИМАТЕЛЬСТВА  
И ИННОВАЦИОННОГО РАЗВИТИЯ  
ГОРОДА МОСКВЫ



РАЗВИТИЕ  
ЧЕЛОВЕЧЕСКОГО  
КАПИТАЛА



ЛИДЕРЫ  
ЦИФРОВОЙ  
ТРАНСФОРМАЦИИ

## Система водоснабжения Мосводоканала:

**Направленный граф** с узлами и ребрами.

**Узлы** – точки измерения потоков и точки соединения линий водоснабжения (параметры – расход, давление, температура, параметры качества воды)

**Ребра** – линии водоснабжения (параметры – материал, длина, диаметр, глубина, коррозионная активность почв, покрытие, катодная защита, дата ввода в эксплуатацию, дата последнего ремонта и т.д.)

**Потребители** – конечные узлы графа, имеют определенные паттерны потребления (ХВС/ГВС)

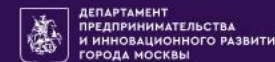
**Цель решения:** нахождение инцидентов, сбор и структурирование данных для предсказания потребления ресурсов, предсказание возникновения прорывов/утечек, в перспективе – обоснование установки частотных приводов на насосы для повышения энергоэффективности и снижения аварийности системы водоснабжения. Также перспективный анализ действий диспетчеров по разрешению инцидентов

## Основные допущения:

- ▶ В идеале: расход воды на входе = сумме расходов на выходе (минус нормативные потери)
- ▶ Каждый потребитель имеет свой паттерн потребления воды в зависимости от часа, дня недели, сезона, рабочих и выходных дней, погодных условий, экономической ситуации и т.д.
- ▶ Инциденты, – отклонение паттернов потребления, превышение допустимой разницы в измерениях, внезапное падение или рост расхода/давления в системе, отсутствие сигнала, внешняя информация.
- ▶ Анализ статистических данных позволит найти ключевые параметры, влияющие на потребление воды, возникновение инцидентов
- ▶ Анализ и кластеризация потребителей может помочь в поиске несанкционированных врезок, утечек, проблем с измерительными устройствами в системе



# ПАЙПЛАЙН РЕШЕНИЯ – 1-й этап (для ИТП)



## Подготовка исходных данных

- Суммирование расходов ХВС для всех ИТП и расходов ГВС для потребителей за каждый час / другой промежуток
- Подготовка прочих данных – погода, экономические данные, праздники и выходные и т.д.

## Поиск и изолирование выбросов, аварий, аномалий

- Сборка датасетов для обучения моделей потребления и анализа инцидентов (сбалансированный датасет)

## Подготовка модели для обучения

- Модели для предсказания потребления и разницы между приборами ГВС/ХВС на N часов вперед на основании данных временных рядов и дополнительных фич (отличие только в target):
- CNN + Transformers + FFNN

## Обучение на данных датасета за весь доступный срок

- Оптимизация модели предсказания потребления и предсказания разнотечения приборов

## Модуль анализа реального потребления vs предсказанного

- Поиск мгновенных значимых расхождений
- Анализ динамики потребления
- Анализ отклонений в показаниях приборов и их динамики

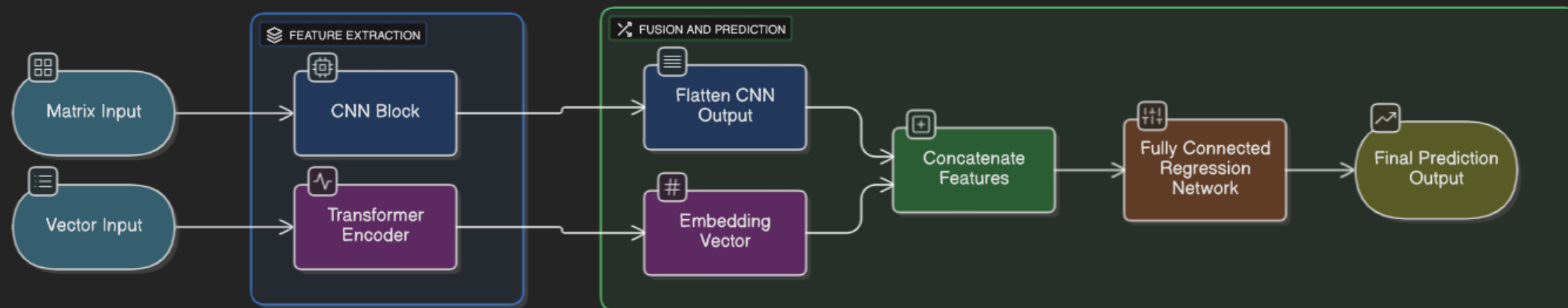
## Построение сервиса по анализу

- Блок визуализации (список ИТП (номер, адрес) с потенциальными инцидентами, подсветка на карте Москвы на основе data.mos.ru)
- Блок подготовки рекомендаций (LLM на основе регламентов/отчетов бригад и диспетчеров и т.д.)
- Чат-бот поддержки диспетчера (RAG системы и локальная LLM)

# АРХИТЕКТУРА МОДЕЛИ (ВАРИАНТ)

- **МОДЕЛЬ ПРЕДСКАЗАНИЯ РАСХОДОВ/ПАТТЕРНОВ ПОТРЕБЛЕНИЯ**
- **Сверточная сеть** – анализ матрицы паттернов за предыдущее время ( $N \times 24$ ), где  $N$  – число дней в ретроспективе
- **Энкодер** – создание эмбединга из вектора всех релевантных параметров (день недели, неделя года, температура, индикаторы выходного/праздничного дня до и после текущего, локация и т.д.)
- **Полносвязный регрессор** – выдача предсказания на  $M$  дней вперед (почасовые данные –  $M \times 24$  часов)
- **Скорость обучения** модели для 10000 ИТП и 2 лет данных на GeForce RTX 3090 – **50 сек/эпоха**

В блок аналитики / сравнения уходят усредненные данные предсказаний пула моделей (разные размеры окна обучения и глубина предсказания).



## 05 – БЛОК АНАЛИЗА

- **Предсказанные ГВС vs предсказанные данные потребления ХВС**
- Ранний поиск признаков проблем
- Анализ систематических ошибок – калибровка?



- **Предсказанные vs реальные данные потребления ГВС/ХВС**
- Анализ накопленного отклонения
- Анализ трендов в разнице (линейные) – текущие средние за 2 разных периода
- XGBoost/RandomForestRegressor

- **Обновление моделей**
- Оптимизация архитектуры
- Если расхождение с реальностью велико
- Наличие новых данных

- **ГВС/ХВС реальные данные за 5-10 мин**
- Отклонение более 10% длительное - сигнал
- Резкий рост/падение до 0 - сигнал

- **Анализ отклонений, анализ ИТП, классификация инцидентов**
- Кластеризация ИТП/Потребителей
- Кластеризация типов отклонений
- Сравнение с результатами оценки специалистами, поиск ошибок
- Scikit-learn K-neighbors etc.

# ЦЕЛИ ПЕРВОГО ЭТАПА



ПРОЕКТ  
МЭРА  
МОСКВЫ



ДЕПАРТАМЕНТ  
ПРЕДПРИНИМАТЕЛЬСТВА  
И ИННОВАЦИОННОГО РАЗВИТИЯ  
ГОРОДА МОСКВЫ



РАЗВИТИЕ  
ЧЕЛОВЕЧЕСКОГО  
КАПИТАЛА



ЛИДЕРЫ  
ЦИФРОВОЙ  
ТРАНСФОРМАЦИИ



- Главная цель – сбор и подготовка данных для обучения модели предсказания технологических ситуаций (инциденты, аварии, реагирование, false positives/negatives, и т.д.)
- Сбор данных регламентов/ отчетов/ интервью для подготовки блока выдачи рекомендаций



- Подготовка синтетических данных, построение и обучение модели – **см. Приложения (файлы)**
- Анализ паттернов потребления моделью, их аномалий и проверка найденных отклонений для выявленных групп ИТП – верификация модели, подтверждение концепции

# ГРАФОВЫЕ СЕТИ ДЛЯ ЭТАПА 2



ПРОЕКТ  
МЭРА  
МОСКВЫ



ДЕПАРТАМЕНТ  
ПРЕДПРИНИМАТЕЛЬСТВА  
И ИННОВАЦИОННОГО РАЗВИТИЯ  
ГОРОДА МОСКВЫ

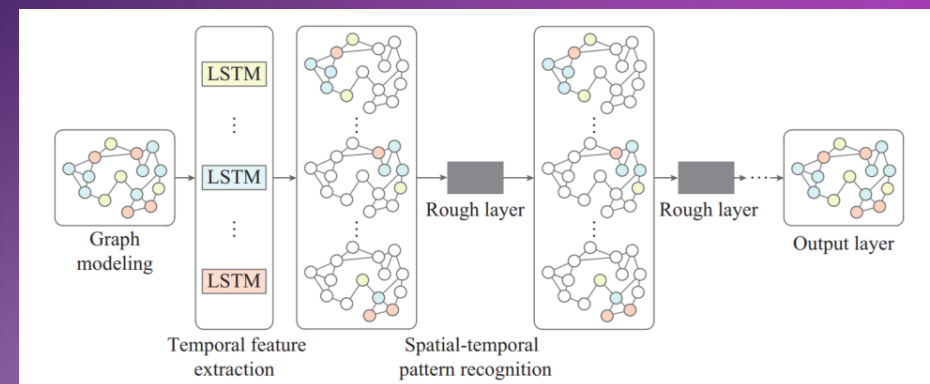
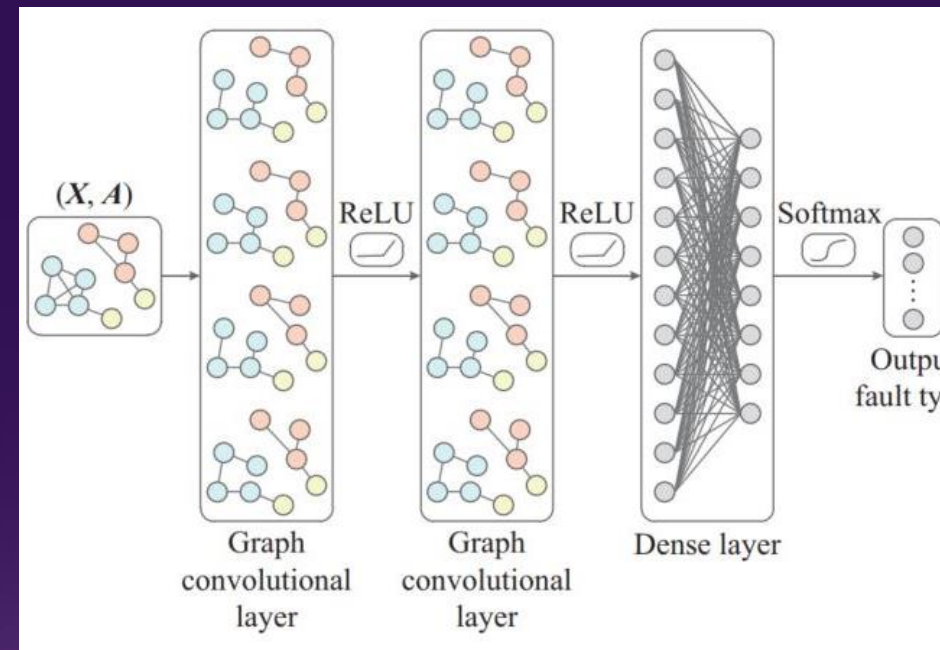
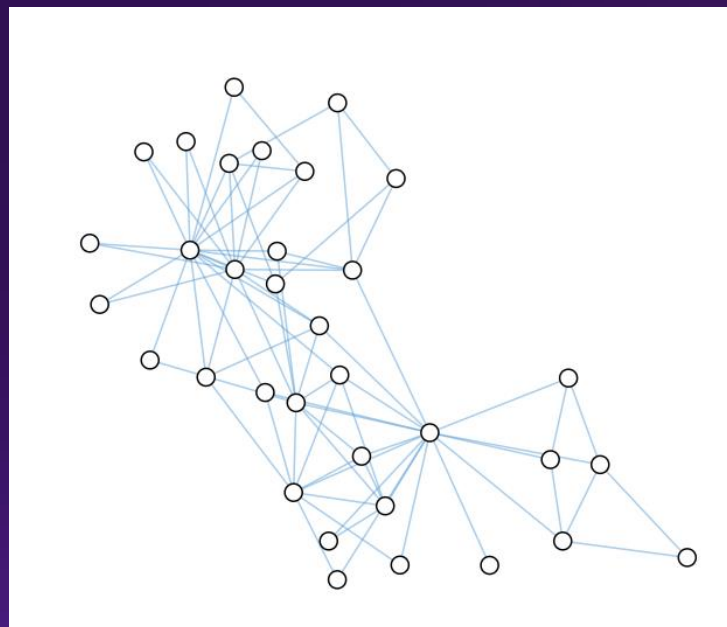
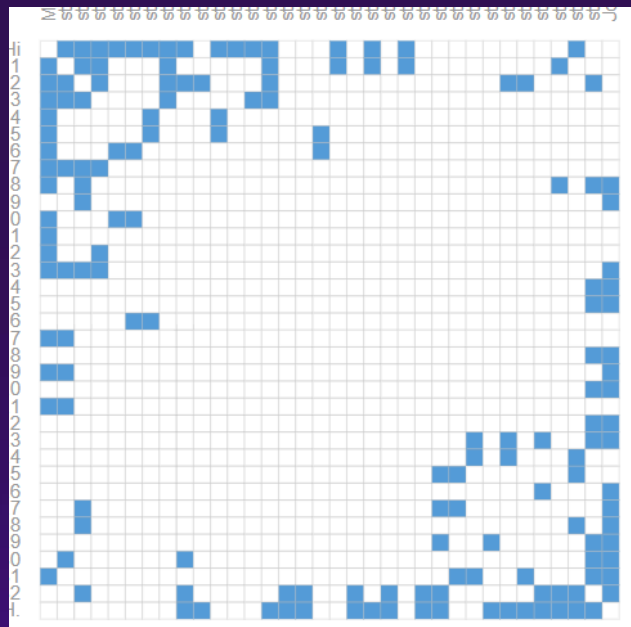


РАЗВИТИЕ  
ЧЕЛОВЕЧЕСКОГО  
КАПИТАЛА



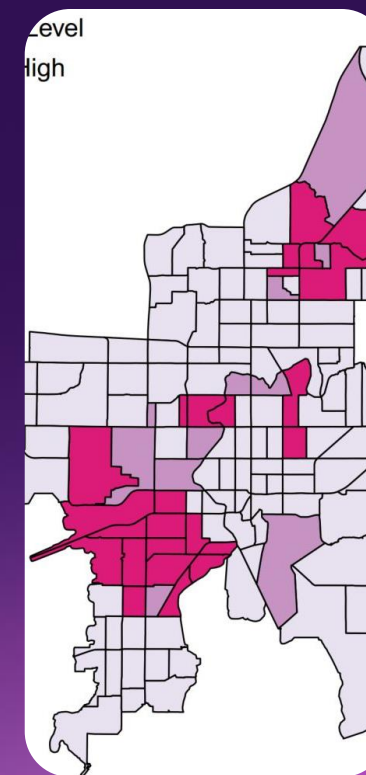
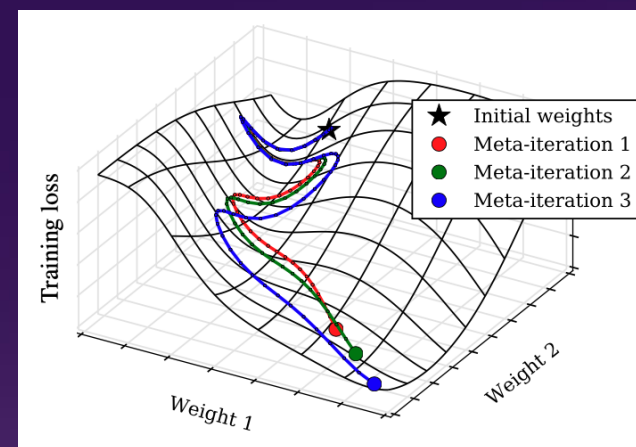
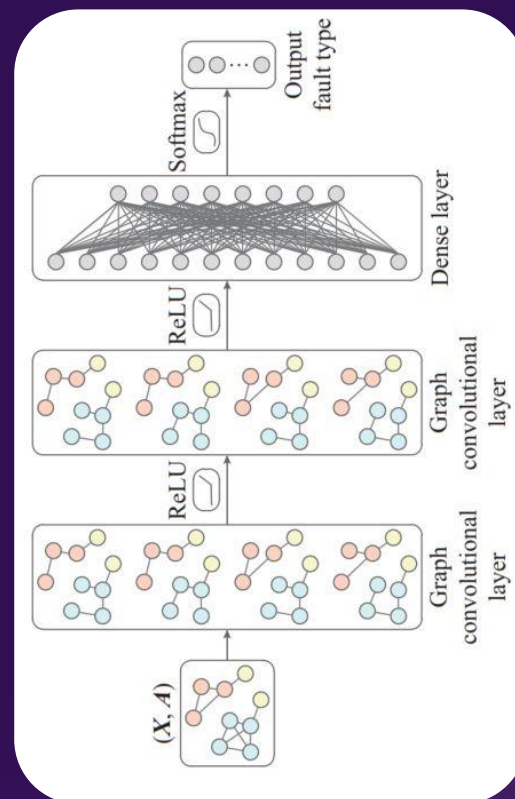
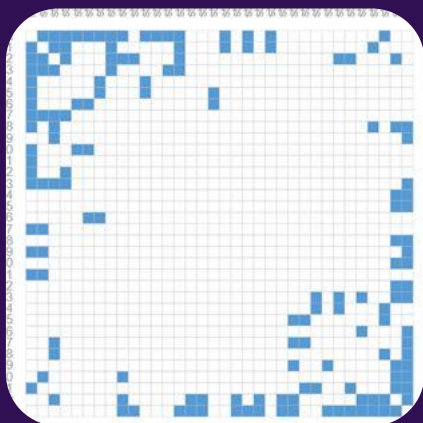
ЛИДЕРЫ  
ЦИФРОВОЙ  
ТРАНСФОРМАЦИИ

- Система Мосводоканала – направленный граф, в котором соседние узлы и ребра влияют на друг друга.
- Узлы – отрезки сети (набор параметров – длина, материал, диаметр, глубина, даты ввода в эксплуатацию и ремонтов и т.д.)
- Ребра – данные замеров счетчиков/датчиков между отрезками (расход, давление, температура)
- Перспективный подход для поиска аномалий и предсказания аварий – использование темпорально-пространственных графовых нейронных сетей для поиска временных зависимостей + сверточные сети (пространственные паттерны)





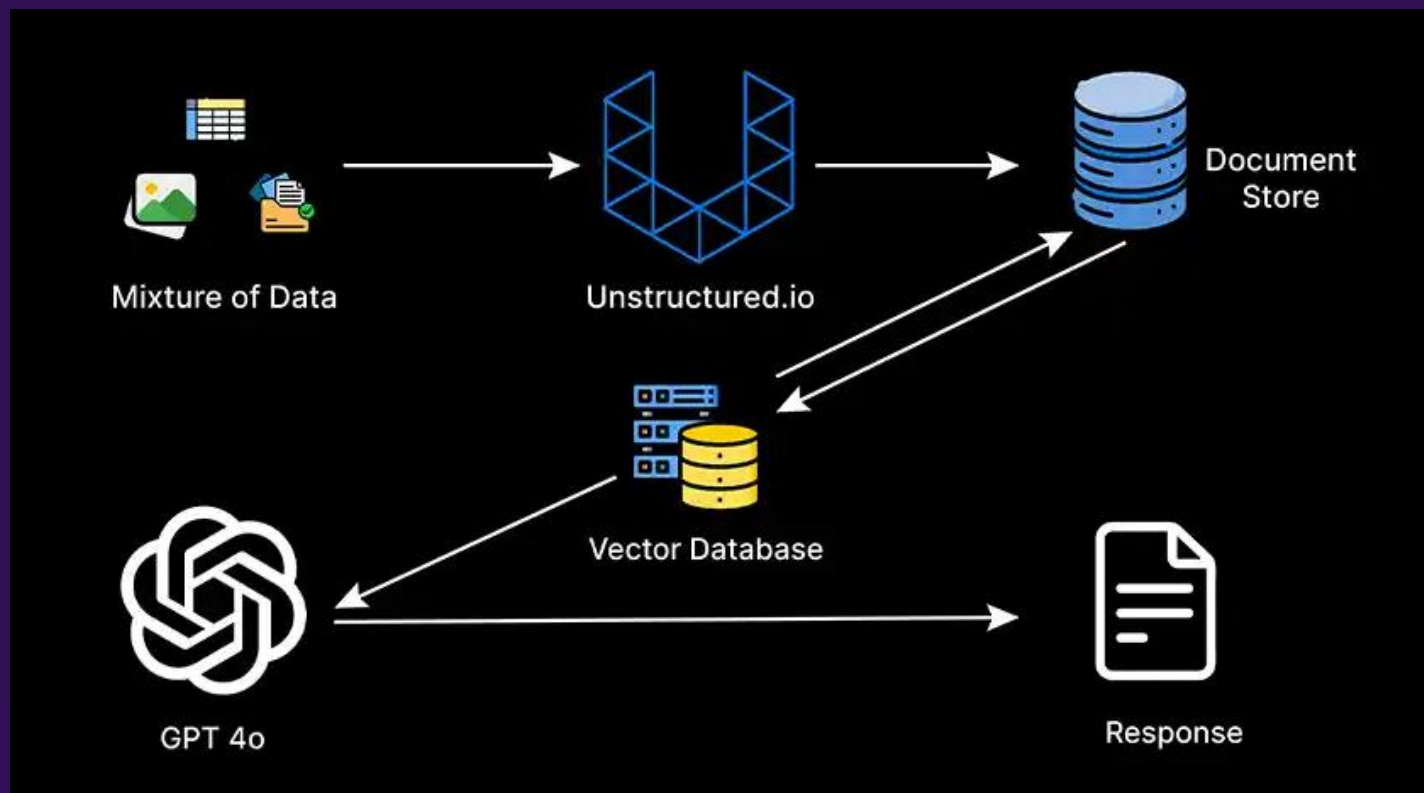
## 2-Й ЭТАП



- Построение графа (матрица связности, параметры узлов и ребер)
- Модель GNN (сверточная, темпорально-пространственная) выбор архитектуры
- Обучение модели для анализа данных, внутренних зависимостей в системе и предсказания технологических ситуаций
- Надстройка и настройка графического интерфейса для удобства руководства, аналитиков и диспетчеров



## 2-Й ЭТАП – LLM и помощники диспетчера



- Построение RAG системы регламентов, отчетов, правил, стандартов, инструкций и т.д.
- Локальная LLM в периметре Мосводоканала как базовый компонент чат-бота – помощника диспетчера
- Объединение модели предсказания с экспертной языковой моделью для выдачи рекомендаций на основе RAG системы – обучение LLM с подкреплением
- Настройка и настройка чат-бота или другого подхода для удобства руководства, аналитиков и диспетчеров

## Этап 3 - Перспективы

### • Мобильное приложение

- Для операторов/специалистов в полях
- Помощь при поиске места аварии, при работе и подготовке отчета

### • Приложение на mosvodokanal.ru

- Сообщить Мосводоканалу о проблеме
- Пришли фото и геолокацию проблемы
- Получить обратную связь



### • Расширение на другие водоканалы

- Есть что улучшать не только в Москве

### • Управление оборудованием

- Изменение давления и подачи в зависимости от прогноза потребления для повышения энергоэффективности и увеличения эксплуатационной готовности
- Планирование профилактики сетей и их расширения



- Сбор данных – сложно в консервативных многоуровневых организациях
- Доступ к отчетам/анализу инцидентов и аварий – важен, но требует ресурсов
- Управление изменениями в организации:
- Наличие четких и понятных целей проекта для всех
- Мотивирование людей, от которых зависит доступ к данным и внедрение системы
- Выделение ресурсов
- Пилот в рамках сегмента сетей – минимизация риска неудачи и оценка экономического эффекта проекта
- Цифровизация – серьезный экономический эффект на сотни млн. руб. в год – абсолютно реально по опыту многих проектов

**Спасибо за внимание!**