

# Разработка системы поиска аномалий во временных рядах температуры воздуха в помещениях

Студенты: Кособанов Денис, Дмитриев Денис

Научный руководитель: Ромазанов Артур Ринатович

# Проблема

В период отопительного сезона важно контролировать температуру помещений. Диапазон допустимых показателей температуры определено в соответствии с требованиями охраны труда. В связи со спецификой каждого отдельного помещения и влияния внешних условий, в случайный период времени, температура воздуха может выходить за допустимые пределы. С помощью собранных данных с датчиков температуры в виде временных рядов, можно выявить возникновение аномалий. Цель проекта - разработать модель нейронной сети для выявления аномалий во временных рядах температур помещений.

# Поиск аномалий во временных рядах зарегистрированных поездок на такси в Нью-Йорке



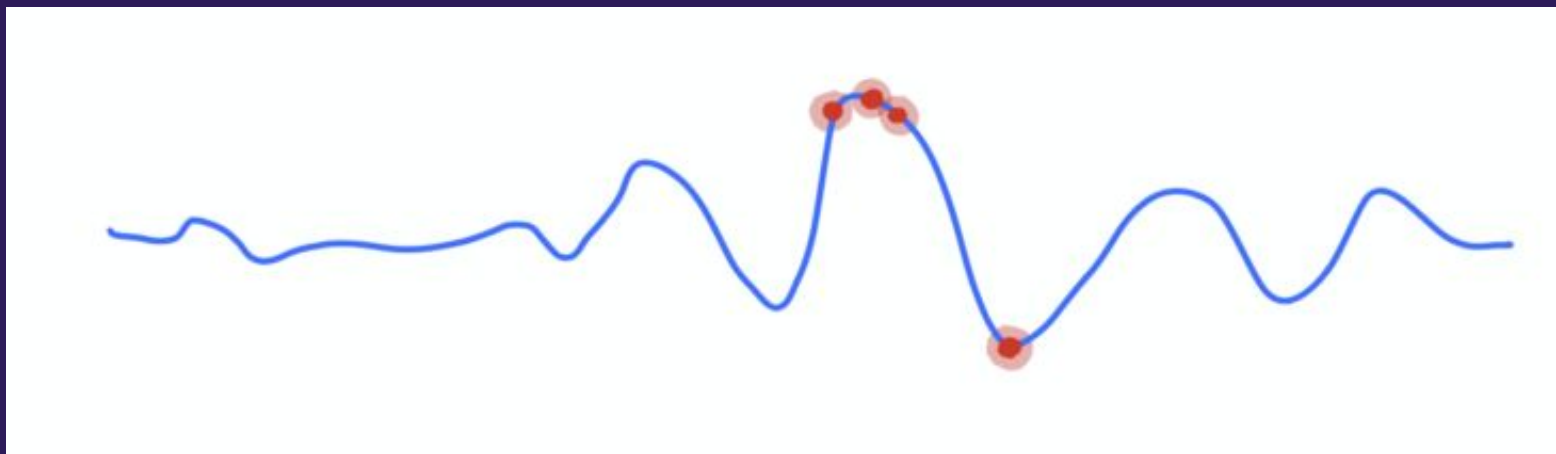
График загруженности такси



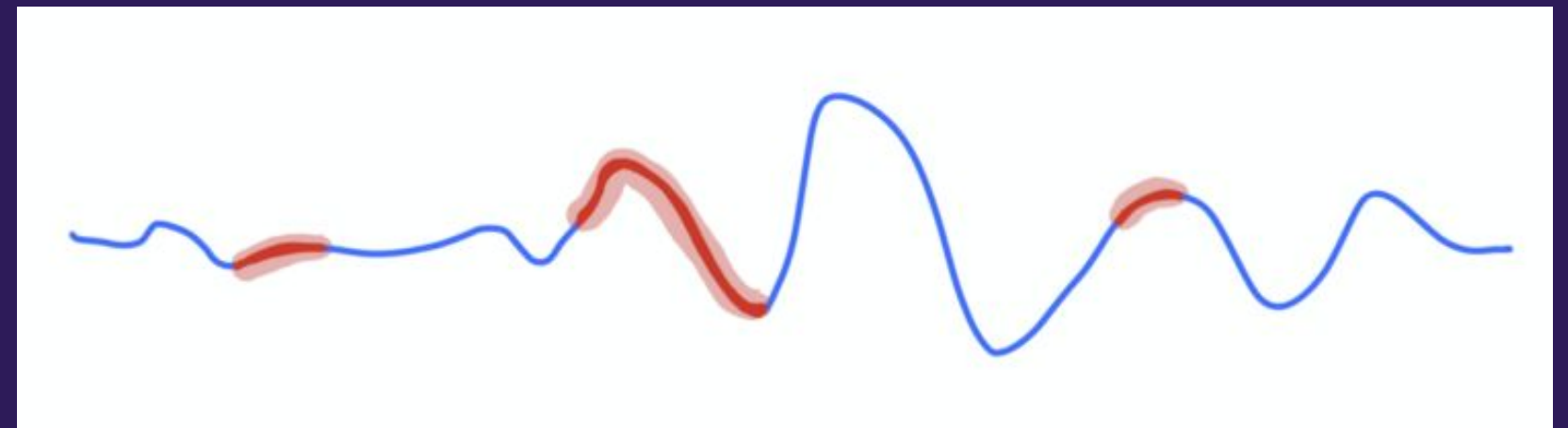
Выявление явных аномалий

# Типы аномалий

- Точечные аномалии - это единичные значения, которые попадают в области значений с низкой плотностью.
- Контекстуальные аномалии - это значения, которые не попадают в регионы с низкой плотностью, но являются аномальными по отношению к локальным значениям.



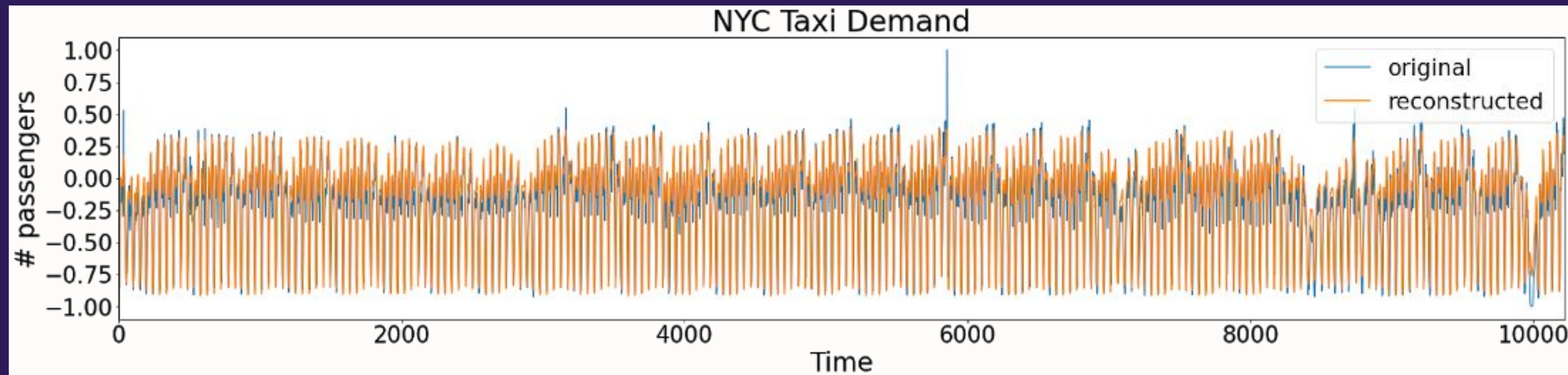
Точечные аномалии



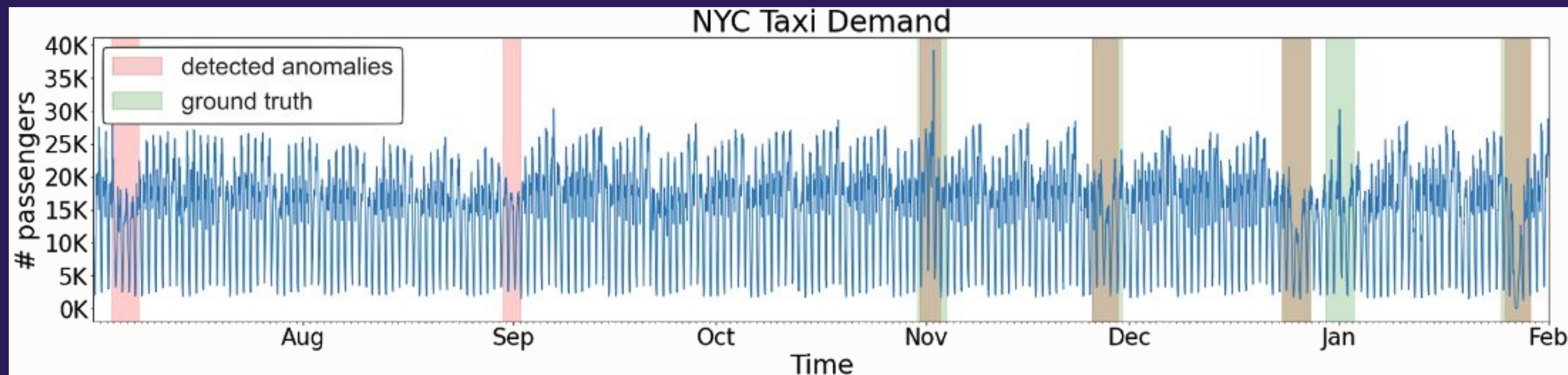
Контекстуальные аномалии

# Использование генеративных состязательных сетей - GAN.

## Результаты.



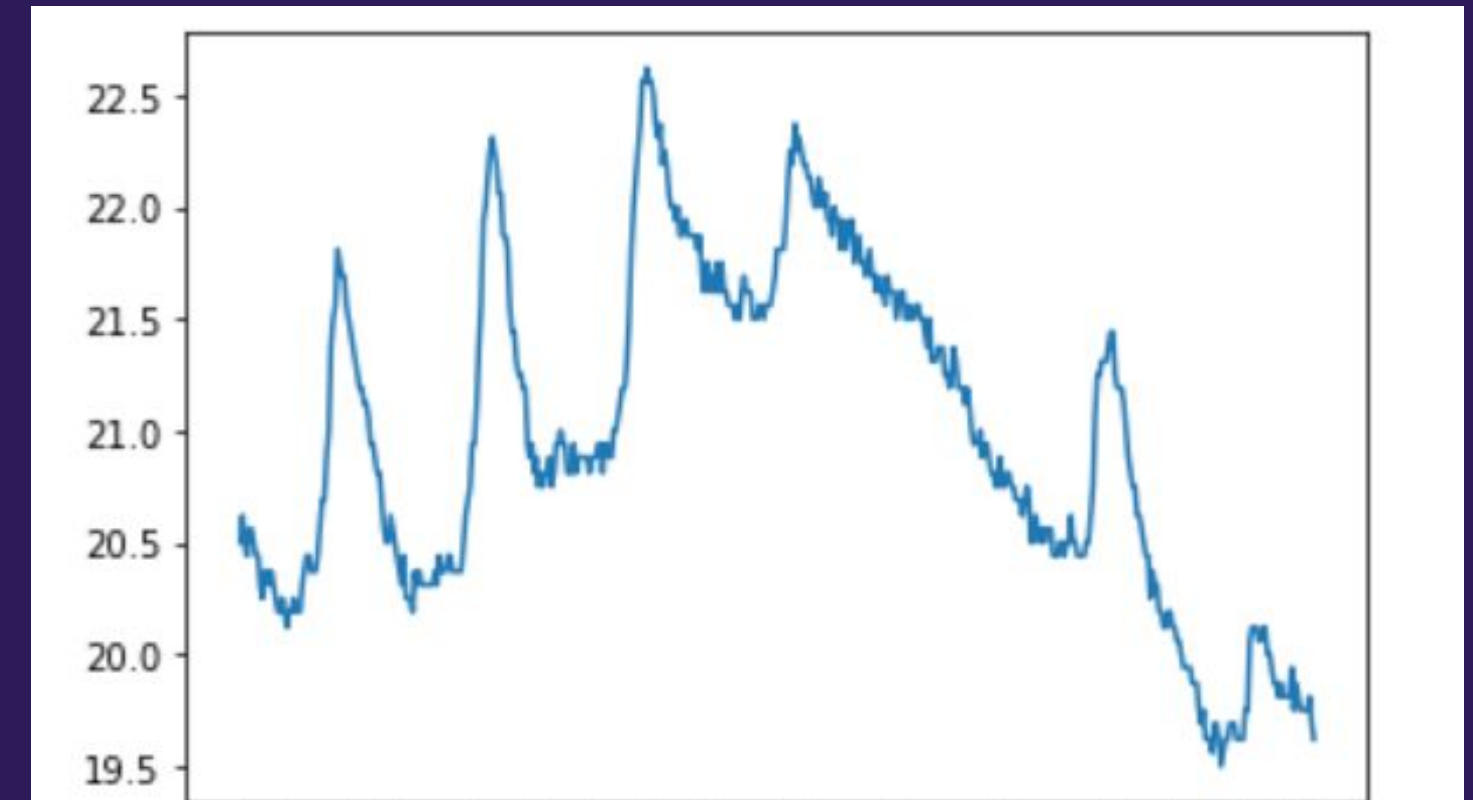
Реконструированный сигнал с использованием GAN, наложенный поверх исходного сигнала



# Предложенное решение

Входные данные могут быть описаны как многомерные временные ряды  $X = x(1), x(2), \dots, x(n)$ , где  $x(t)$  принадлежит одномерному пространству  $R^m$ ,  $n$  – число точек времени.

- Поиск аномалий временного ряда аудитории. reconstruction-based методы
- Поиск аномальных аудиторий. Для обнаружения аномальных аудиторий будут использоваться метрические алгоритмы.



Ненормализованный график температуры за первую неделю 2021 года

## Предполагаемые инструменты и технологии. TadGAN

Одним из недавно разработанных reconstruction-based методов, показывающих хорошие результаты в обнаружении аномалий, является TadGAN, разработанный исследователями из MIT в конце 2020 года. Архитектура метода TadGAN содержит в себе элементы автокодировщика и генеративных состязательных сетей.

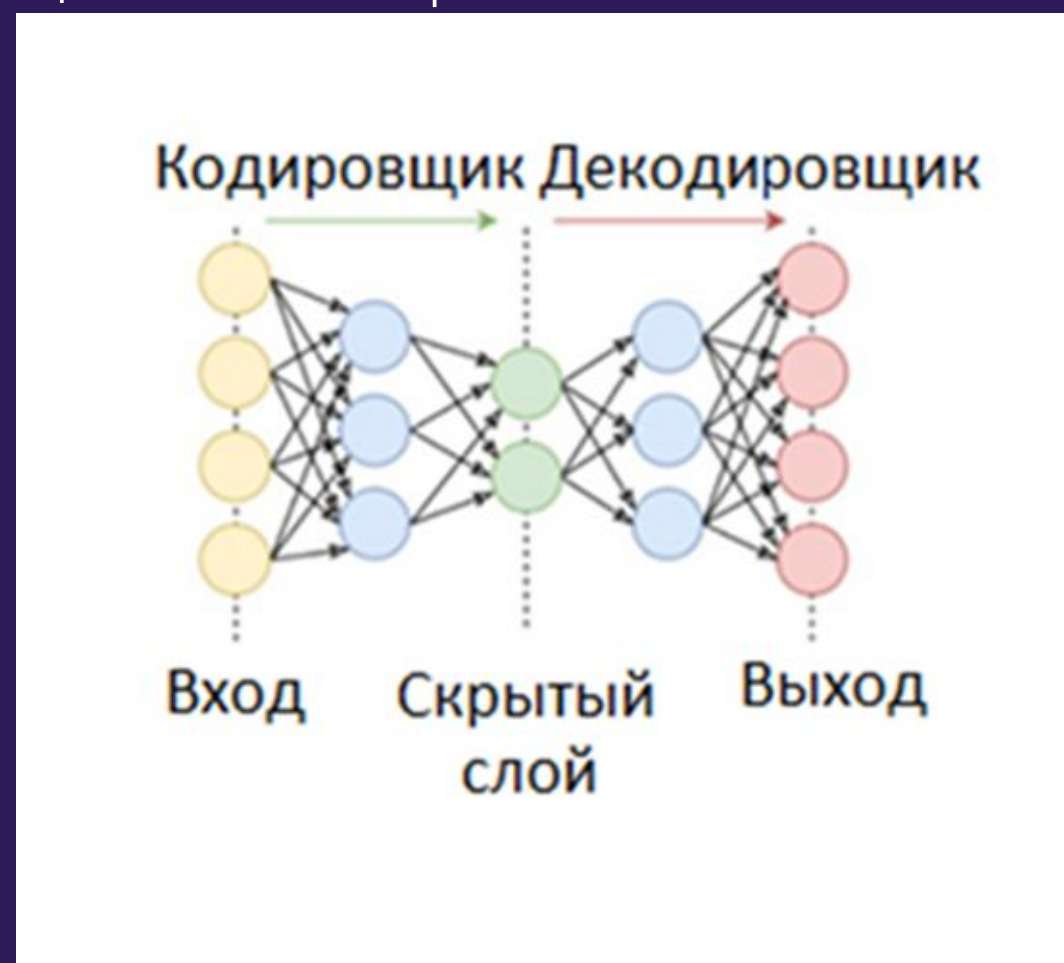
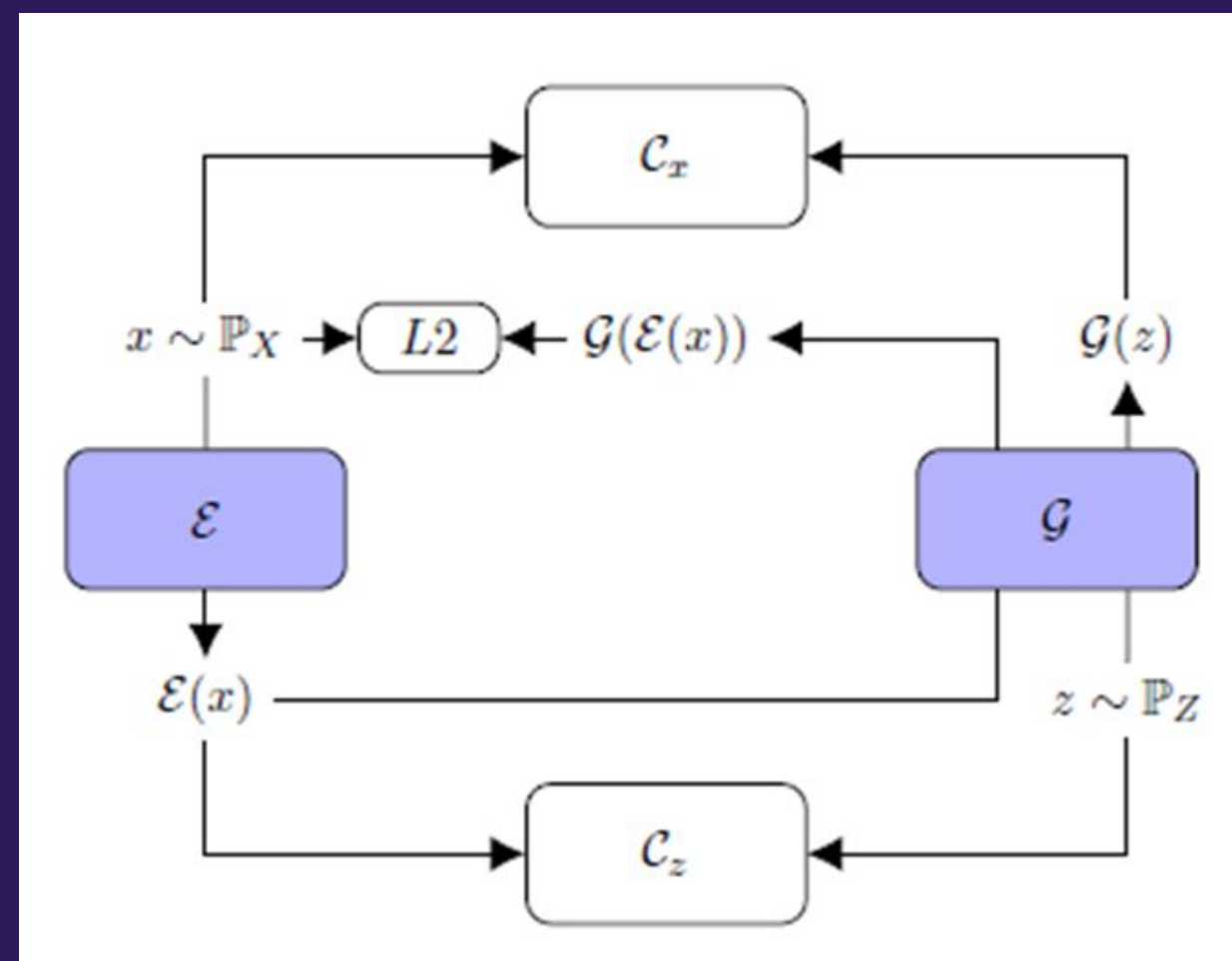


Схема работы кодировщика



Архитектура метода TadGAN

## Предполагаемые инструменты и технологии. TadGAN

Для создания и обучения нейронной сети могут быть использованы различные стандартные пакеты (например, TensorFlow или PyTorch), имеющие высокоуровневое API. При обучении модели оптимизировались следующие пять метрик:

- aeLoss
- cxLoss
- cx\_g\_Loss
- czLoss
- cz\_g\_Loss

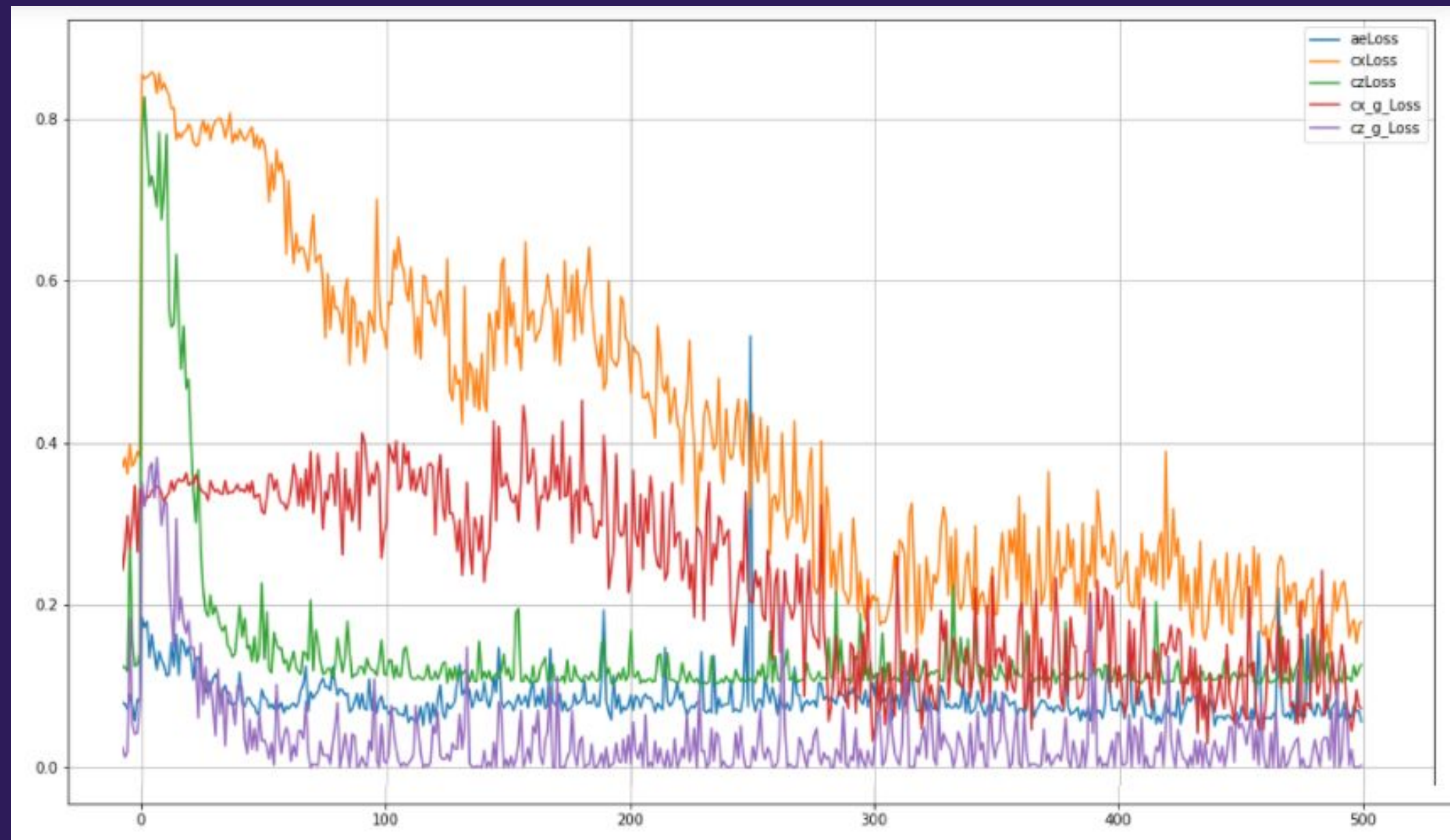


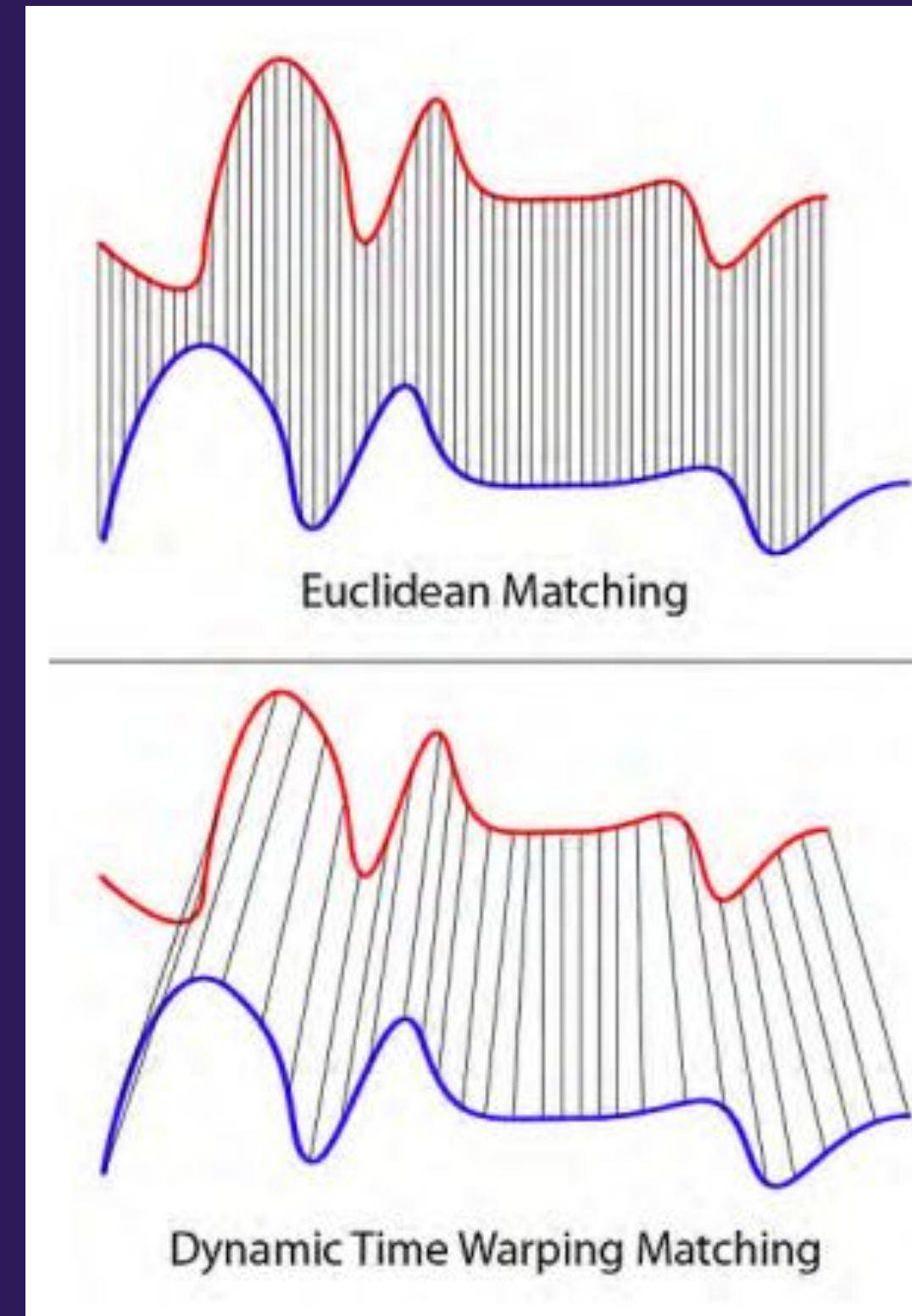
График обучения модели TadGAN для 500 эпох



# Предполагаемые инструменты и технологии. Метрические методы

k-NN: расстояние от временного ряда до его k-го ближайшего соседа в наборе данных временных рядов - это показатель аномалии.

Несогласованность фаз и нелинейные выравнивания различных временных рядов, которые являются некоторыми общими проблемами для данных временных рядов, ограничивают использование различных мер приближения для этих классов методов.



Отличие DTW-евклидовой метрики от евклидовой

# Инструменты оценки результата. Алгоритма TadGAN

|                          |   |   |
|--------------------------|---|---|
|                          | Аномалия предсказана моделью,<br>$y_{\hat{}} = 1$ | Модель предсказала отсутствие аномалии  |
| Аномалия есть<br>$y = 1$ | TP<br>правильно предсказанная аномалия            | FN<br>аномалия есть, но ее не нашли     |
| Аномалии нет<br>$y = 0$  | FP<br>предсказали аномалию там, где ее нет        | TN<br>аномалии нет и модель ее не видит |