



Факультет компьютерных наук

ОП Прикладная математика и
информатика

Москва 2024

Анализ ЭКГ на основе ML

ECG Analysis on the ML Algorithms

Исследовательский проект

Выполнила: Зыкова-Мызина Анна Кирилловна, 229 группа, 2 курс

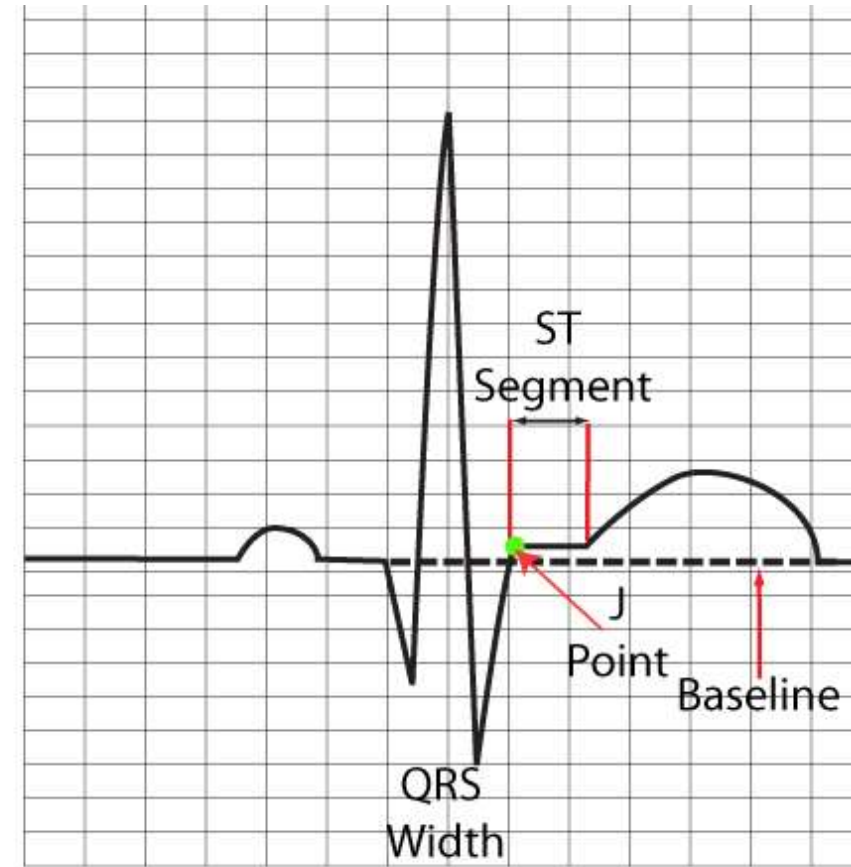
Научный руководитель: Хельвас Александр Валериевич
старший преподаватель, МФТИ

Описание предметной области

Анализировалось одно отведение **24-часовых записей холтера**

Предметом анализа являлось **смещение ST-сегмента** – отклонение изоэлектрической линии вверх или вниз, возникающее после QRS-комплекса

Обычно его измеряют через 60-80 миллисекунд после J-точки





Актуальность работы

Ручной анализ больших объемов данных холтеровского мониторинга сложен и склонен к ошибкам, а автоматизация позволит врачам быстрее и точнее выявлять потенциально опасные аномалии.

Актуальность данной работы заключается в отсутствии достаточного числа работ, касающихся детекции аномалий **ST** сегмента в данных холтера, а также потенциальной важности проблемы.



Цель и задачи работы

Цель работы – исследовать эффективность альтернативных подходов поиска аномальных эпизодов смещения **ST**-сегмента в записях холтера по одному отведению.

Задачи работы:

1. Написать функции для расчета ЧСС и построения спектрограммы по записи холтера
2. Изучить соответствующие исследования, связанные с детекцией аномалий **ST** сегмента
3. Написать функцию, которая бы разделяла **24**-часовую запись холтера на отдельные удары
4. Написать функцию которая измеряла бы степень отклонения **ST** сегмента и обучить модель **LSTM** автоэнкодер для поиска аномалий в одномерном временном ряду ($\Delta ST \text{ segment} / time$)
5. Обучить модель **CNN** для бинарной классификации ударов сердца (аномальный / нормальный)
6. Оценить качество двух построенных моделей



Анализ существующих решений

1. В работе **Interpretable Assessment of ST-Segment Deviation in ECG Time Series [1]** решалась задача многоклассовой классификации:
 - сегментация более длительных записей ЭКГ (одно отведение) на отдельные удары сердца
 - вычленение некоторых признаков (длины сегментов, медианные и средние значения всего сигнала...)
 - дальнейшее обучение векторов признаков (с помощью **General Automated Machine learning Assistant**, а также **LSTM** и **CNN**)
2. В работе **Deep Learning Networks Accurately Detect ST-Segment Elevation Myocardial Infarction and Culprit Vessel [2]** решалась задача бинарной классификации (10 секундных размеченных участков ЭКГ для 12 отведений):
 - были предложены 3 нейронной сети (**CNN**, **LSTM** и **LSTM – CNN**)



Описание датасета

Индивидуальные записи **базы данных Long-Term ST** имеют продолжительность от **21** до **24** часов и содержат два или три сигнала ЭКГ.

Также приложена разметка: записаны интервалы эпизодов, во время которых изменение величины **ST** – сегмента превосходило **$V_{min} = 75$ мкВ** и длилось не менее **$T_{min} = 30$ с**



LSTM автокодировщик для поиска аномалий в одномерном временном ряду (ΔST segment / time)

24-часовая запись
холтера



Разделение записи на
отдельные удары
сердца



Вычисление смещения
ST-сегмента для каждого
удара



Составление временного
ряда ΔST segment /
time



Поиск аномалий с
помощью модели LSTM
autoencoder

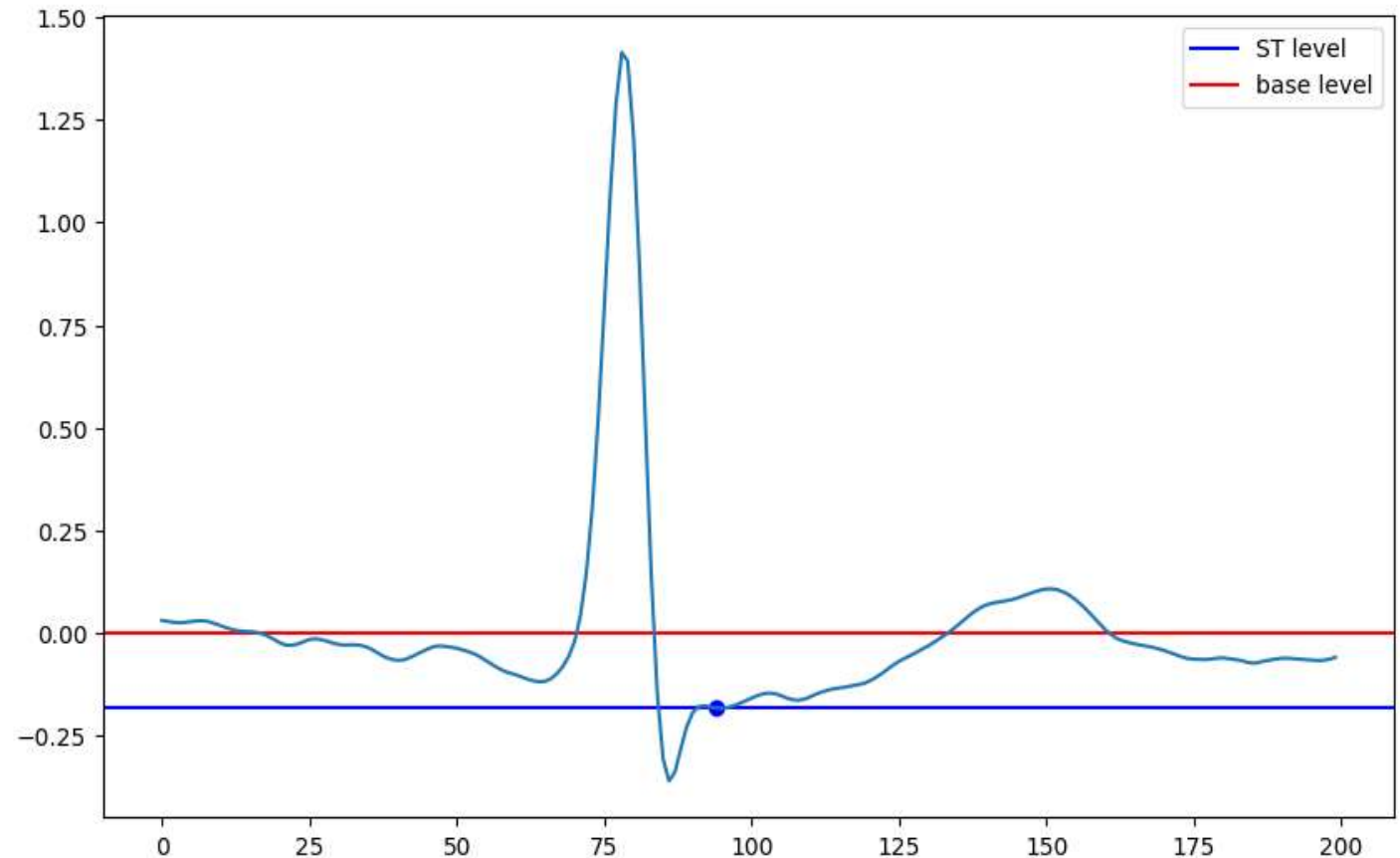


Оценка эффективности
модели



Смещение ST сегмента вычислялось на основе алгоритма, предложенном в статье **Wireless Single-Lead versus Standard 12-Lead ECG, for ST-Segment Deviation during Adenosine Cardiac Stress Scintigraphy**[3]

1. Выравнивание изоэлектрической линии
2. Определение точки с наименьшей производной на RT интервале





1 этап: Подготовка данных

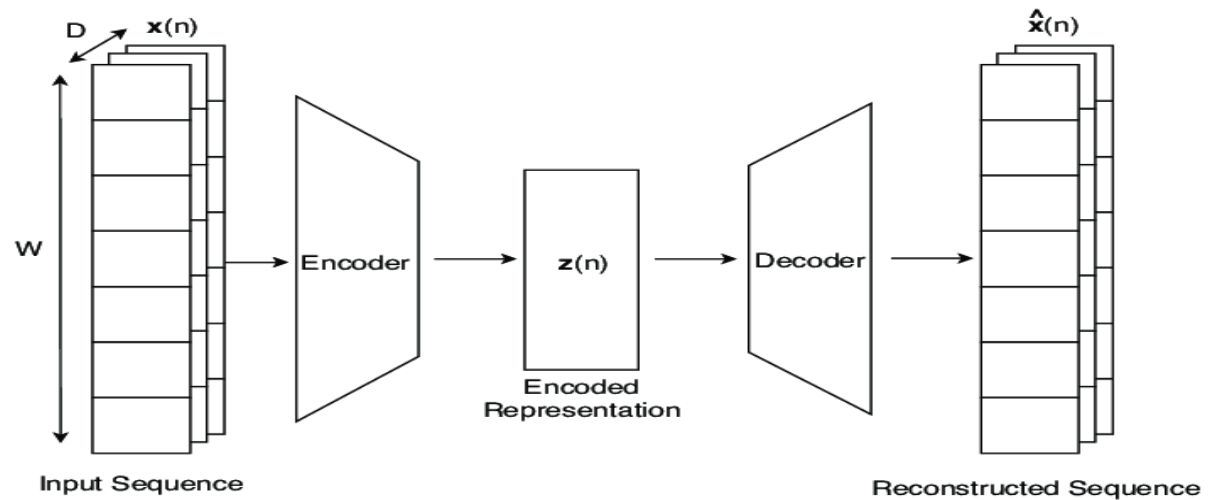
Удаление размеченных аномалий из
тренировочных данных

Использование скользящего среднего для
уменьшения шума

Min-Max нормализация

Создание трёхмерного тензора с помощью окон
с наложением для train и test

2 этап: Обучение на train



LSTM(128)
LSTM(64)
Dropout

Repeated Vector

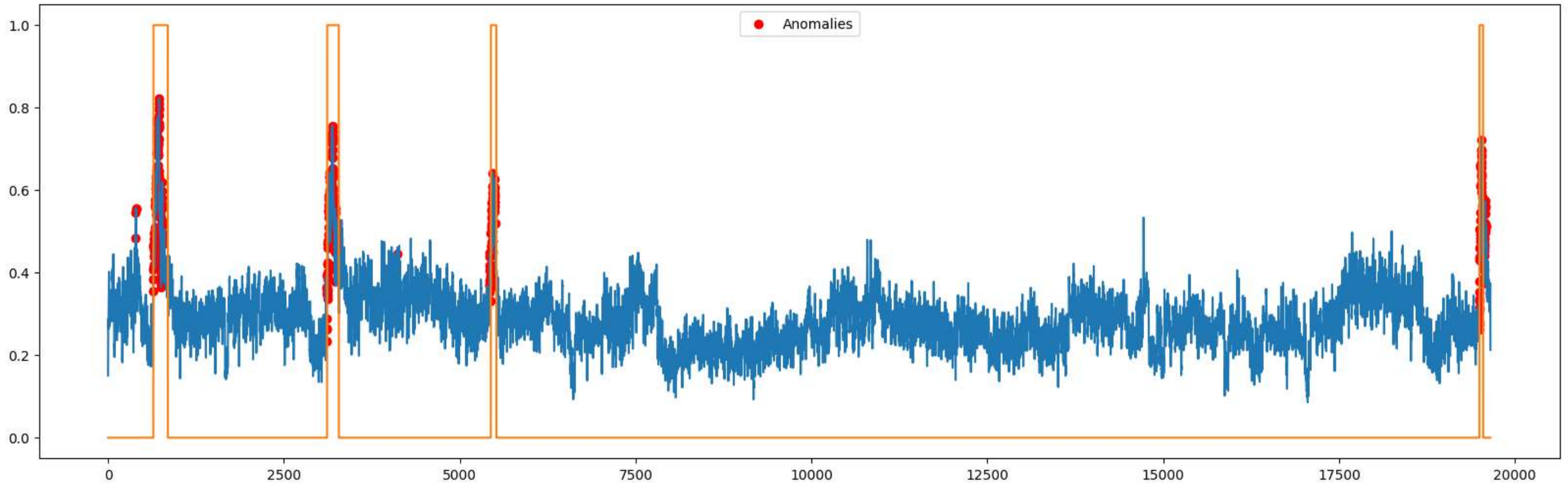
LSTM(64)
LSTM(128)
Dropout



3 этап: обнаружение аномалий на test

Расчет ошибки предсказания для каждого интервала

Интервал x помечался аномальным, если $\text{err}(x) > \mu + 2,5\sigma$

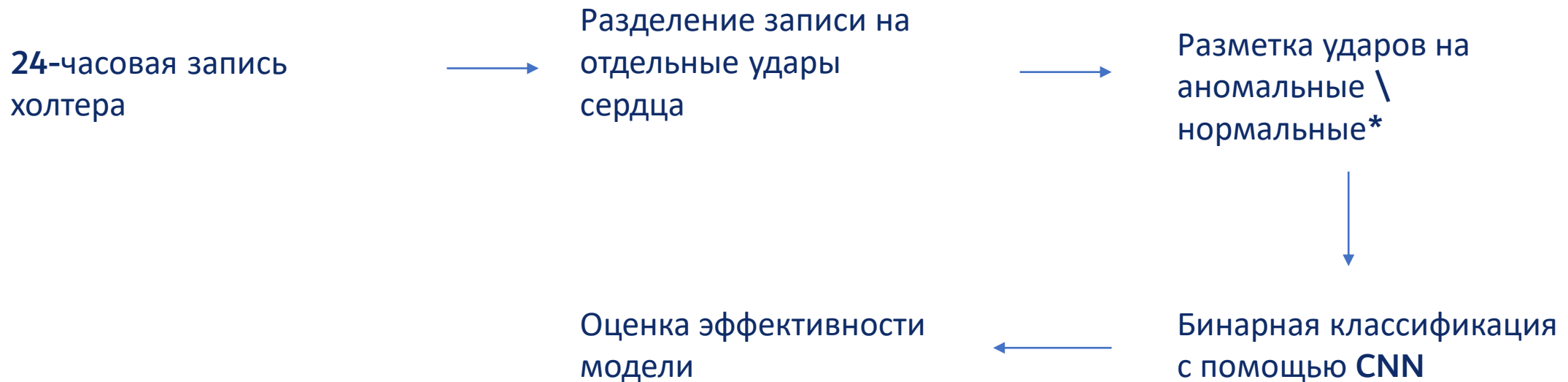




Пациент	True Positive	False Positive	False Negative	Precision	Recall
1	4	2	0	0.67	1.00
2	5	3	1	0.62	0.83
3	5	4	1	0.56	0.83
4	5	5	2	0.50	0.71
5	2	1	2	0.67	0.50
6	6	2	2	0.75	0.75
7	7	4	3	0.64	0.70
8	3	1	1	0.75	0.75
9	4	6	1	0.40	0.80
10	16	5	5	0.76	0.76
Всего	57	33	18	0.64	0.76



Свёрточная нейронная сеть (CNN) для бинарной классификации аномальных и нормальных ударов



* Удар помечался как аномальный, если он попадал в размеченный отрезок (величина смещения $>75\text{мВ}$ и продолжительность $>30\text{ сек}$), что является неточной разметкой



1 этап: подготовка данных

Были выделены отдельные удары сердца, начиная с пика S и заканчивая пиком T

Min – Max нормализация

156993 нормальных и 71936 ударов сердца (обработано 30 холтеров)

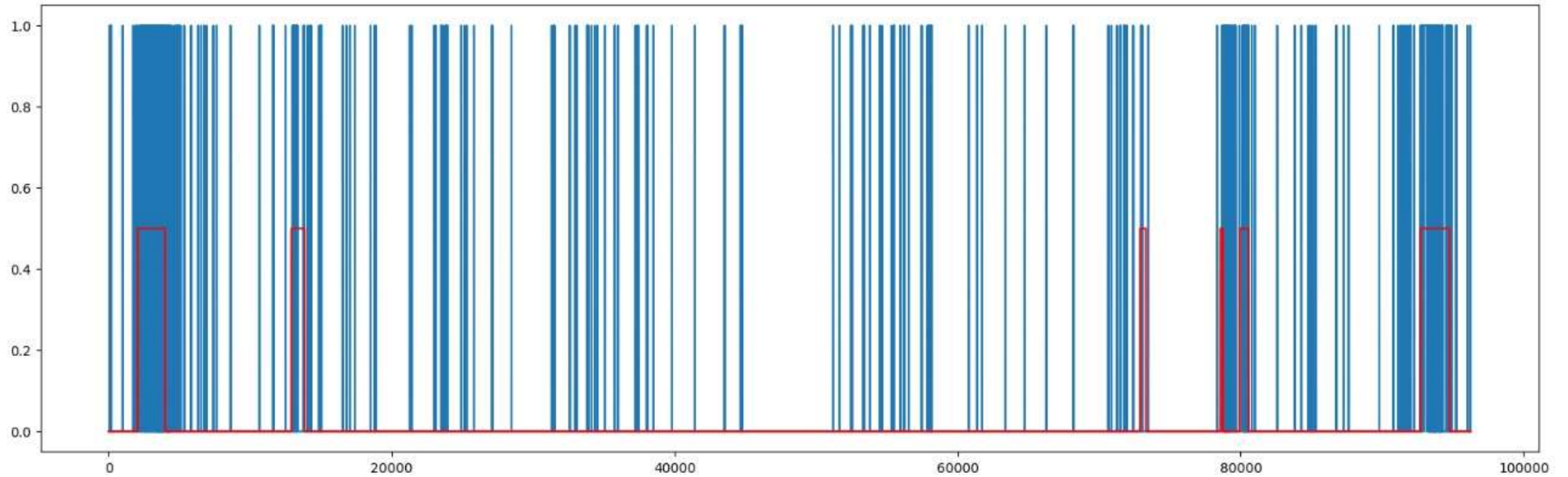
2 этап: обучение модели - CNN сеть, принимающая на вход одномерный массив и возвращающая значения 0 / 1

3 свёрточных слоя с различным количеством фильтров: 32, 32, 64 с функцией активации **relu**, а также слои **Batch Normalization**
После 2 и 3 свёрточных слоёв **MaxPooling1D** с размером окна 2

3 этап: тестирование

Accuracy - 0.90, Precision - 0.29, Recall - 0.44, F1-Score - 0.35

Класс	Precision	Recall	F1-Score	Support
0	0.96	0.93	0.94	90,099
1	0.287	0.437	0.347	6,142
Accuracy				0.894
Macro avg	0.62	0.68	0.64	96,241
Weighted avg	0.92	0.89	0.90	96,241





Результаты

Написаны функции для первичной обработки данных холтера, а также 2 модели, решающие задачу поиска аномальных эпизодов смещения **ST** – сегмента.

1. LSTM автоэнкодер продемонстрировал свою эффективность: метрики составили **Precision – 0.63** и **Recall – 0.76**.

Перспективы работ: написать более точную функцию для измерения **ST** – сегмента.

2. CNN продемонстрировал такие показатели как: **Accuracy - 0.90, Precision - 0.29, Recall - 0.44, F1-Score - 0.35**, что говорит о низкой эффективности использования данного подхода на базе данных **Long-Term ST**.

Перспективы работ: протестировать предложенную модель на более тщательно размеченных данных, где каждый сердечный удар отмечен как нормальный или аномальный.



СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1) Interpretable Assessment of ST-Segment Deviation in ECG Time Series / Israel Campero Jurado, Andrejs Fedjajevs, Joaquin Vanschoren, Aarnout Brombacher // Sensors. — 2022. — Vol. 22, no. 13. <https://www.mdpi.com/1424-8220/22/13/4919>
- 2) Deep Learning Networks Accurately Detect ST-Segment Elevation Myocardial Infarction and Culprit Vessel / Lin wu, Guifang Huang, Xianguan Yu et al. // Frontiers in Cardiovascular Medicine. — 2022. — 03. — Vol. 9.
- 3) 6. Wireless Single-Lead versus Standard 12-Lead ECG, for ST-Segment Deviation during Adenosine Cardiac Stress Scintigraphy / Luna Fabricius Ekenberg, Dan Eik Høfsten, Søren M. Rasmussen et al. // Sensors. — 2023. — Vol. 23, no. 6. <https://www.mdpi.com/1424-8220/23/6/2962>.

