

Пояснительная записка

к итоговому проекту на тему:

**«Классификация заболевания по форме ЭКГ»**

Автор: Ферапонтов Сергей

Группа: DLL-28

Оглавление

1. Постановка задачи для глубокого обучения
2. Анализ данных
3. Методика реализации
4. Итоги обучения модели
5. Выводы и заключение

# Постановка задачи

## Введение в электрокардиографию

Электрокардиограммы (ЭКГ) являются распространённым методом диагностирования заболеваний у пациентов. Аппарат измеряет разность потенциалов в 12 или 24 отведениях, производя таким образом график 12 или 24 величин, синхронизированных по времени [1]. Различные аппараты проводят измерения с различной частотой и точностью, поэтому на практике может потребоваться небольшая предобработка данных. Современные медицинские стандарты не предписывают учёта никаких дополнительных факторов (пол, возраст, вес и т.д.) при постановке диагноза. Отметим, что так как финальное решение всегда принимает лечащий врач, разрабатываемый алгоритм следует рассматривать как систему поддержки принятия решений, призванную автоматизировать работу младшего медицинского персонала, расшифровывающего ЭКГ.

Особенностью данного метода является широкая доступность машинного представления показаний прибора, что отражается в наличии крупных датасетов с легко интерпретируемыми данными.

Ручная обработка сигнала ЭКГ содержит два этапа:

* + - 1. Измерение зубцов и интервалов;
      2. По полученным значением вынесение диагноза по таблице.

Таким образом, автоматизировать ЭКГ-диагностику с применением машинного обучения возможно, анализируя исходный временной ряд показаний прибора или же анализируя измененные величины зубцов и интервалов.

Остановимся на первом методе, так как для решения второго лучше подойдут алгоритмы, основанные на деревьях решений и им подобные. Кроме того, потенциальный алгоритм анализа временного ряда позволит избавиться от необходимости установки модуля измерений зубцов и интервалов.

Отметим, что измеренные прибором данные могут быть представлены как в виде временного ряда, так и в виде фотографии плёнки ЭКГ. Для работы используем именно представление в виде временного ряда, что более приближенно к автоматической обработке показаний аппарата ЭКГ.

Также важной особенностью данного типа данных является связь один ко многим между исходными данными и диагнозами (Multilabel), что затрудняет применение ряда базовых моделей.

### Описание выбранного датасета

Исходя из описанных выше ограничений, выберем датасет «PTB-XL ECG dataset [2]». Для разработки модели используем данные в виде последовательностей с частотой измерений 100 Гц.

Каждому примеру соответствует не менее одной метки, для каждой метки в примере дана степень уверенности. Всего ра

Ниже приведём основные характеристики используемых данных.

# Список литературы

1. Электрокардиография: учебн. пособие / В.В.Мурашко, А.В.Струтынский. – 18-е изд. – Москва : МЕДпреcс-информ, 2022. – 360 с. : ил.