**Постановка задачи**

<https://archive.ics.uci.edu/ml/datasets/Heart+Disease>  
Измерение параметров состояния человека при болезни сердца.

Структура:  
age - Возраст  
sex - Пол: 0 - женский, 1 - мужской  
cp - Тип грудной боли: 1 - типичная ангинальная, 2 - атипичная ангинальная, 3 - неангинальная, 4 - безсимптомная  
trestbps - Кровяное давление  
chol - Концентрация холестерола  
fbs - Уровень сахара в крови  
restecg - Результат кардиограммы: 0 - нормальный, 1 - ST-T аномалия, 2 - гипертрофия левого желудочка  
thalach - Максимальный пульс  
exang - Боль вызвана физическими упражнениями: 0 - нет, 1 - да  
oldpeak - Уменьшение ST-сегмента после физических упражнений  
slope - Форма ST-сегмента после физических упражнений: 1 - наклон вверх, 2 - плоский, 3 - наклон вниз  
ca - Количество окрашенных крупных сосудов при флюороскопии  
num - Диагноз стеноза артерии: 0 - сужение диаметра менее 50%, 1 - сужение диаметра более 50%

Задача: вычисление значения параметра age на основании остальных параметров.

**Подход к обработке данных.**

В столбцах age, chol, thalach была использована функция StandardScaler().fit\_transform(data).flatten() во всех остальных столбцах использовали MinMaxScaler().

К столбцам age, thalach был применен логарифм так как они давали график похожий на нормальное распределение или были близки к нему. В остальных столбцах использовался квадратный корень. Квадратный корень давал лучший график или присутствовали нули.

В столбцах Trestbps, Chol, Oldpeack было большое количество нулей в нашем случае это говорит об отсутствие данных, поэтому мы заменили нули на среднее значение.

Sex, exang, restecg, num, fbs, ca, cp –В данных столбцах были тестовые значения которые привели к числовым группам.

Trestbps, Chol, Thalach, Oldpeack были выбросы к которым был применен клипинг.

**Базовые параметры нейронной сети.**

функция активации - ReLU

Количество нейронов входного слоя – 13

Первый скрытый слой – 15

Второй скрытый слой – 15

Выходной слой – 1

Количество эпох – 50

Батчей – 16

**Описание проведенных экспериментов.**

1 модель - сделал модель с 2 слоями и небольшим количеством нейронов. Графики не очень плавные, но модель работает корректно.  
2 модель –Изменил параметры, сначала добавил 3 слой график расходился, увеличил количество эпох, точность стала выше, количество необходимых параметров стало больше.

Итоговые параметры 2 модели.

функция активации - ReLU

Количество нейронов входного слоя – 13

Первый скрытый слой – 5

Второй скрытый слой – 8

Третий скрытый слой - 15

Выходной слой – 1

Количество эпох – 75

Батчей – 16  
3 модель – Попробовал оставить 1 скрытый слой. Точно стала хуже, графики почти не изменились, параметров требует еще больше чем 2 модель.

Итоговые параметры 3 модели.

функция активации - ReLU

Количество нейронов входного слоя – 13

Первый скрытый слой – 50

Выходной слой – 1

Количество эпох – 50

Батчей – 16

**Итоговый вывод.**

Увеличение скрытых слоев нейронной сети делает алгоритм нейронной сети более сложным. Время работы сети также увеличивается с количеством слоев. Если у нас малый набор данных и большое количество слоев мы можем переобучить нейронную сеть или сделать ее решение очень специфичным. Чем больше слоев, тем больше нейронной сети требуется параметров.