**1. Постановка задачи**

<https://archive.ics.uci.edu/ml/datasets/QSAR+aquatic+toxicity>  
Информация о структуре молекулы вещества и его токсичности для определенного вида рыб.

Структура:  
tpsa - Площадь поверхности молекулы, с атомами, которые могут соединяться с фосфором и серой  
saacc - Площадь поверхности молекулы, с атомами, которые могут соединяться с водородом  
h-050 - Количество атомов, которые могут соединяться с водородом  
mlogp - Коэффициент липофильноси (стремления модекулы соединяться с ограническими веществами)  
rdchi - Топологический коэффициент (определяет форму молекулы)  
gats1p - Коэффициент поляризуемости молекулы  
nn - Количество атомов азота  
c-040 - Количество атомов углерода в определенных связях  
lc50 - Концентрация, при которой погибает более 50% рыб

Задача: вычисление значения параметра lc50 на основании остальных параметров.

**2. Подход к обработке данных.**

Ко всем столбцам был применен квадратный корень так как он давал результат лучше логарифма.

Во всех столбцах кроме 'mlogp', 'rdchi', 'gats1p', 'lc50' заменил нули на среднее значение.

Применил StandardScaler() к 'mlogp', 'rdchi', 'gats1p', 'lc50' так как их графики ближе к нормальному распределению.

Выбросы были во всех столбцах.

**3.Базовые параметры нейронной сети.**

функция активации - ReLU

Количество нейронов входного слоя – 7

Первый скрытый слой – 6

Второй скрытый слой – 6

Выходной слой – 1

Количество эпох – 50

Батчей – 8

**4. Описание проведенных экспериментов.**

Первая модель работает корректно и показывает нормальные графики.

Во второй модели использовал функцию Softmax. Для первого и второго слоя установил количество нейронов 8. Графики сошлись и модель работает корректно.

В третьей модели использовал функцию sigmoid. Для первого и второго скрытого слоя установил количество нейронов равным 12. и также сошлись.

**5. Итоговый вывод.**

Relu требует большого количество нейронов, функция очень медленная.

Функция не линейна, и функция представляет из себя горизонтальную линию, некоторые нейроны могут из-за этого не работать.

Третья модель лучше реагирует на “шероховатости” данных. У краев функции реакции функции на изменение “X” слабее, это значит, что некоторые предсказания могут быть занижены.

Softmax не требует большого количества нейронов, функция устойчивая из-за чего ее часто используют в связке с Relu. Также softmax считается улучшенной версией линейной функции.