**Постановка задачи**

Источник: https://opendata.mkrf.ru/opendata/7705851331-stat\_museum

Данные о музеях России: размер коллекции, график работы и т.п.

Структура:

name - Название музея

base\_fond - Число предметов основного фонда на конец года

science\_fond - Число предметов научно-вспомогательного фонда на конец года

ams - Автоматизированная музейная система: число внесенных музейных предметов

n\_buildings - Число строений

n\_days - Число дней в году, открытых для посещения

guests - Число посещений выставок, экспозиций + экскурсионных посещений, тысяч человек

employees - Численность работников, человек

Задача: вычисление значения параметра guests на основании остальных параметров.

**Подход к обработке данных.**

Ко всем столбцам кроме base\_fond , science\_fond , guests было применено MinMaxScaler().fit\_transform(data).flatten().

Выбросы были в столбцах science\_fond, ams, n\_days, n\_buildings, Employees, Guests к этим столбцам применилnp.clip(data, 0, 0).

В столбцах science\_fond, n\_buildings, guest, employees было большое количество нулей и отсутствующие значения. Я заменил их на среднее значение.

Квадратный корень применил к столбцам ams, n\_buildings, emloyees.

Name – не стал его обрабатывать так как в нем уникальные и независимые значение.

**Базовые параметры нейронной сети.**

функция активации - ReLU

Количество нейронов входного слоя – 6

Первый скрытый слой – 6

Второй скрытый слой – 6

Выходной слой – 1

Количество эпох – 50

Батчей – 16

**Описание проведенных экспериментов.**

Первая модель работала хорошо. Во второй модели я увеличил количество эпох до 75, нейронов потребовалось больше поэтому в первом слое 30 нейронов, а во втором слое 20 нейронов. Графики были хорошие точность почти не изменилась. В третей модели я снизил количество эпох до 25, в первом слое 40 нейронов во втором слое 18 нейронов. График ухудшился точность почти не изменилась.

**Итоговый вывод.**

Произошла одна эпоха (epoch) — весь датасет прошел через нейронную сеть в прямом и обратном направлении только один раз. Чем больше, тем сильнее обучается нейронная сеть. Если эпох слишком много сеть может переобучиться и наоборот если слишком мало эпох мы можем не закрепить правила и недоучить нейронную сеть.