**Постановка задачи**

<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 гожу, 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.**

В столбце base\_fond была использована функция StandardScaler().fit\_transform(data).flatten() во всех остальных столбцах использовали MinMaxScaler().

К столбцам base\_fond, n\_days, employees 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.

Во всех столбцах, кроме base\_fond, были выбросы к которым был применен клипинг.

Столбец name содержит неповторяющиеся имена, поэтому он не использовался.

**Базовые параметры нейронной сети.**

функция активации - ReLU

Количество нейронов входного слоя – 6

Первый скрытый слой – 20

Второй скрытый слой – 20

Выходной слой – 1

Количество эпох – 50

Размер батча – 15

Оптимизатор – adam

**Описание проведенных экспериментов.**

Первая модель работает правильно.

Вторая модель имеет следующие параметры:

Количество нейронов входного слоя – 6

Первый скрытый слой – 20

Второй скрытый слой – 20

Выходной слой – 1

Количество эпох – 50

Размер батча – 100

Оптимизатор - adam

При увеличении размера батча обучение произошло быстрее, распределение стало более «нормальным», то есть ближе к центру гистограммы. Так же, графики динамики отклонений быстрее вышли на стабильный уровень. Графики квантиль-квантиль стали более «прямыми».

Третья модель имеет следующие параметры:

Количество нейронов входного слоя – 6

Первый скрытый слой – 20

Второй скрытый слой – 20

Выходной слой – 1

Количество эпох – 50

Размер батча – 5

Оптимизатор - adam

Обучение происходит дольше, среднеквадратичные ошибки увеличились, графики динамики отклонения выглядят менее «ровно», из чего можно сделать вывод, что на стабильный уровень они не вышли.

Четвертая модель имеет следующие параметры:

Количество нейронов входного слоя – 6

Первый скрытый слой – 20

Второй скрытый слой – 20

Выходной слой – 1

Количество эпох – 50

Размер батча – 500

Оптимизатор - adam

При слишком большой размере батча обучение происходит быстро, но в разы хуже, точность намного меньше. Графики динамики отклонений выходят на стабильный уровень намного позднее.

**Итоговый вывод.**

Размер батча влияет как на скорость, так и на качество обучения. При увеличении размера батча скорость обучения возрастает, однако при слишком большом значении качество может сильно упасть, что видно на примере 4 эксперимента. Самый лучший результат показал эксперимент 2 с размером батча 100, так как получилось наилучшее соотношение скорости и качества обучения.