

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

Факультет систем управления и робототехники

Прикладной искусственный интеллект

Лабораторная работа № 2

Выполнил студент

Шиняков Артём Дмитриевич

Группа № R32372

**Преподаватели:**

Евстафьев Олег Александрович

г. Санкт-Петербург

2022

# Содержание

## Цель

Необходимо по имеющимся данным о ценах на жильё предсказать окончательную цену каждого дома с учетом характеристик домов с использованием нейронной сети. Описание набора данных содержит 80 классов (набор переменных) классификации оценки типа жилья, и находится в файле `data_description.txt`.

В работе требуется дополнить раздел «Моделирование» в подразделе «Построение и обучение модели» создать и инициализировать последовательную модель нейронной сети с помощью фреймворков тренировки нейронных сетей как: Torch или Tensorflow. Скомпилировать нейронную сеть выбрав функцию потерь и оптимизатор соответственно. Оценить точность полученных результатов. Вывести предсказанные данные о продаже.

## Результат (ссылка на google colab):

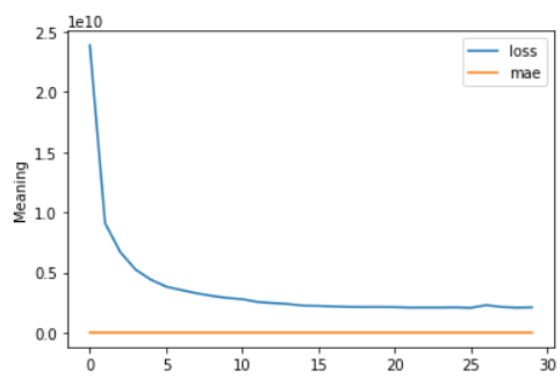
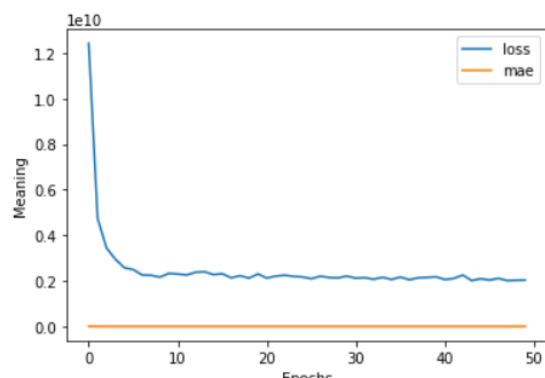
[https://colab.research.google.com/drive/1fuUQyBG7IardZdOOn\\_phRWsW5rMxWtlc?usp=sharing](https://colab.research.google.com/drive/1fuUQyBG7IardZdOOn_phRWsW5rMxWtlc?usp=sharing)

## Ход выполнения:

В ходе подбора и инициализации модели мною было выявлено пять наиболее влияющих на ее работу составляющих:

1. Количество нейронов на входном слое;
2. Количество эпох;
3. Размер мини-выборки;
4. Оптимизаторы;
5. Функции потерь;

Далее я решил создать 6 моделей нейронной сети, чтобы сравнить влияние гиперпараметров и строения модели. Например: количество нейронов на входном слое влияет на строение всей сети, соответственно на результат и нагрузку на вычислительную машину. Количество эпох также влияет на нагрузку, а также на веса внутри модели. Но слишком малое количество эпох может привести к недообучению, а слишком большое к переобучению. При этом по графикам обучения можно довольно четко определить необходимое количество эпох, поскольку в прекрасный момент веса почти перестают меняться.



В итоге по графикам видно, что 30 эпох имеют место быть для минимизации ошибки, но большее количество не показывает улучшений в минимизации для функции потерь. Так же как и количество нейронов в слое больше 200 не давало ничего кроме увеличенной нагрузки, поэтому я остановился на 200 нейронах в слое(несмотря на то, что одинаковая модель, в которой различалось только количество нейронов в скрытом слое показала примерно одинаковый результат) для точности и 50 эпохах(просто для наглядности бессмысленности лишних эпох).

Также большое влияние оказывает выбор функции потерь, то есть наиболее подходящей под нейронную сеть. Например MSE и Huber показали приличные результаты при других гиперпараметрах. А вот сеть с категориальной кроссэнтропией предскаывает цену как отрицательную, что, разумеется, является абсурдом.

```
[1459 rows x 2 columns]
   Id  SalePrice
0  1461  140726.218750
1  1462   93941.679688
2  1463  184590.828125
3  1464  188747.828125
4  1465  165329.578125
...  ...
1454 2915   79950.320312
1455 2916  101805.914062
1456 2917  182806.937500
1457 2918   87325.351562
1458 2919  224181.703125
```

```
[1459 rows x 2 columns]
   Id  SalePrice
0  1461  -457.702820
1  1462   56.027954
2  1463 -468.490234
3  1464 -439.646118
4  1465 -496.941223
...  ...
1454 2915  -361.101776
1455 2916 -295.029663
1456 2917 -492.936279
1457 2918 -476.863220
1458 2919 -331.781067
```

Остальные гиперпараметры всё же не так сильно влияют на результаты модели. Например оптимизатор не повлиял на проблему предсказания кроссэнтропией отрицательной цены.

В итоге лучше всего себя показал MSE, независимо от других параметров. Huber также является хорошим выбором для последовательной модели, а вот Кроссэнтропия плохо подошла под задачу предсказания.

## Ответы на вопросы:

- 1) Как вышеперечисленные параметры влияют на полученный вами результат? - Описано выше
- 2) Что такое эпоха (Epoch)? В чем отличие от итерации (Iteration)? – Эпоха – проход данных вперед и назад по нейронной сети, то есть это цикл полного набора обучающих данных, а итерации - это количество пакетов или шагов через разделенные пакеты обучающих данных, необходимых для завершения одной эпохи.
- 3) Что такое функция активации? Какие вам известны? – Это функция, которая определяет выходной сигнал нейрона сети, опираясь при этом на входной сигнал (или на набор таковых). Знаю функцию единичной ступеньки, арктангенса, тождественную. Однако большинство из них устроены так, чтобы возвращать число от 0 до 1, что удобно при определении весов нейронов, поэтому множество функций похожи друг на друга.
- 4) Что такое MSE (Mean Squared Error) - Средняя квадратичная ошибка? Что такое MAE (Mean Absolute Error)? Для чего используются. – MAE - это мера ошибок между парными наблюдениями, выражающими одно и то же явление. Примеры  $Y$  по сравнению с  $X$  включают сравнения прогнозируемого и наблюдаемого, последующего времени и начального времени, а также один метод измерения по сравнению с альтернативным методом измерения. Средней квадратичной ошибкой называется среднее квадратичное значение из суммы квадратов ошибок отдельных измерений. MAE - статистика, среднеквадратичная ошибка, а MSE – функция риска, которая измеряет среднее значение квадратов ошибок — то есть среднюю квадратическую разницу между оценочными значениями и фактическим значением.