

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра МО ЭВМ

ОТЧЕТ
по лабораторной работе №3
по дисциплине «Искусственные нейронные сети»
Тема: Регрессионная модель изменения цен на дома в Бостоне

Студент гр. 8383

Сахаров В.М.

Преподаватель

Жангиров Т.Р.

Санкт-Петербург

2021

Цель работы.

Реализовать предсказание медианной цены на дома в пригороде Бостона в середине 1970-х по таким данным, как уровень преступности, ставка местного имущественного налога и т. д.

Данный набор содержит относительно немного образцов данных: всего 506, разбитых на 404 обучающих и 102 контрольных образца. И каждый признак во входных данных (например, уровень преступности) имеет свой масштаб. Например, некоторые признаки являются пропорциями и имеют значения между 0 и 1, другие — между 1 и 12 и т. д.

Задачи.

1. Ознакомиться с задачей регрессии
2. Изучить отличие задачи регрессии от задачи классификации
3. Создать модель
4. Настроить параметры обучения
5. Обучить и оценить модели
6. Ознакомиться с перекрестной проверкой

Ход работы.

Была построена модель из методических материалов, последний слой в которой не имеет функции активации, таким образом сеть сможет предсказывать значения из любого диапазона.

Так как задача регрессии отличается от задачи классификации, в модели используется другая функция потерь (mean square error), а также функция mean absolute error на этапе обучения

Исходные данные, из-за их малого количества, были разбиты на 4 блока. Обучение поочередно производилось на 3-х блоках, оставшийся использовался

для проверки. Для нахождения графиков средней ошибки использовалась функция average библиотеки numpy.

Программа со 100 эпохами обучения показала среднюю ошибку, равную 2.57. По графику на рисунке 1 видно, что переобучение началось примерно после 40 эпох.

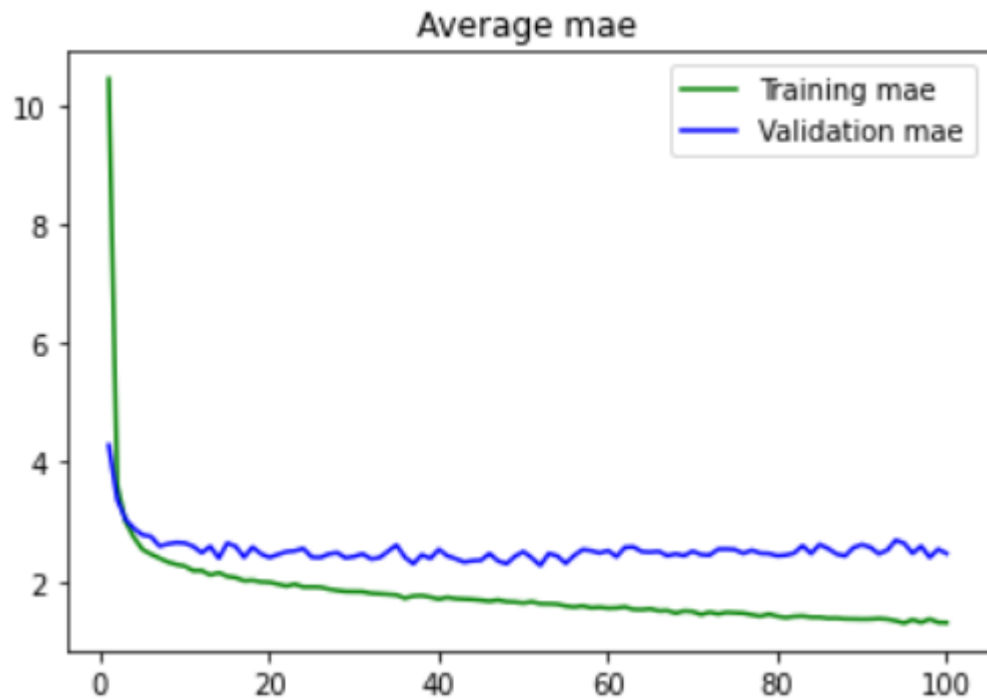


Рисунок 1

Опустив количество эпох до 40, получаем среднюю ошибку, равную 2.37. График изменения средней ошибки показан на рисунке 2.

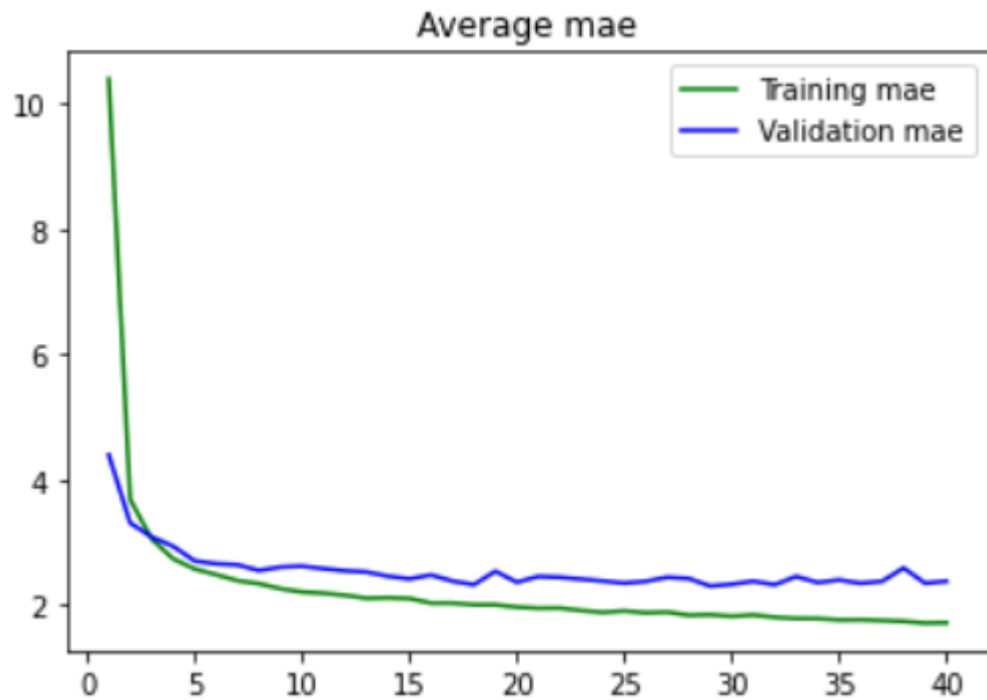


Рисунок 2

Также было проверен результат при увеличении количества блоков до 6 и 8. Средние результаты оказались равны 2.34 и 2.40 соответственно. Графики ошибок приведены на рисунках 3 и 4. Графики ошибок не имеют выраженной точки минимума с низким значением ошибки, так что ощутимого снижения процента ошибок добиться увеличением количества блоков на таком маленьком датасете не представляется возможным

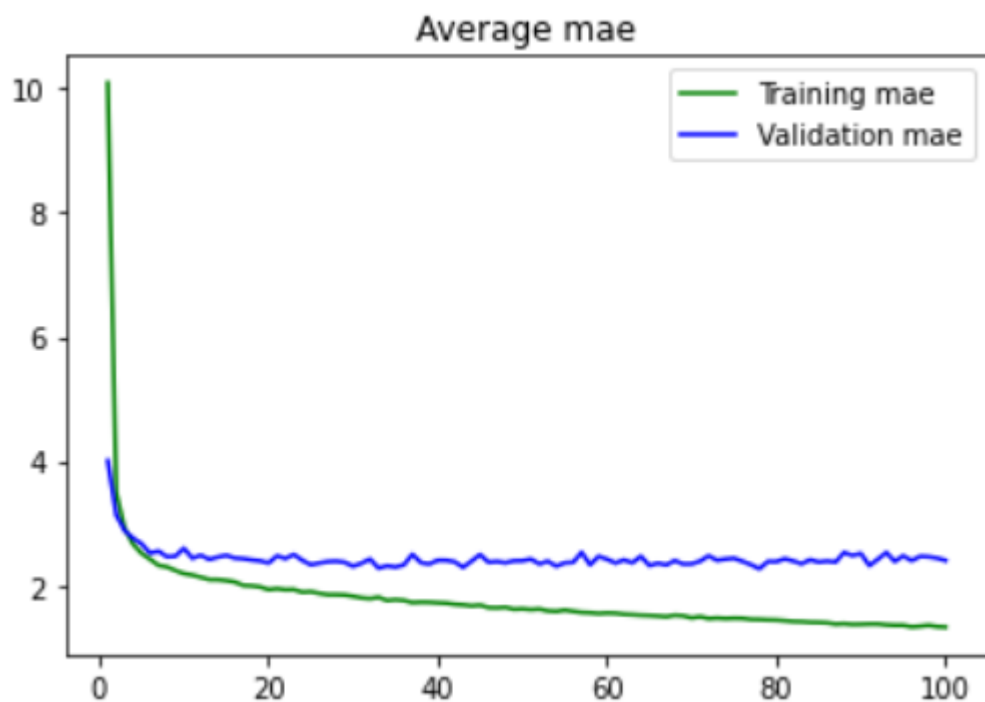


Рисунок 3

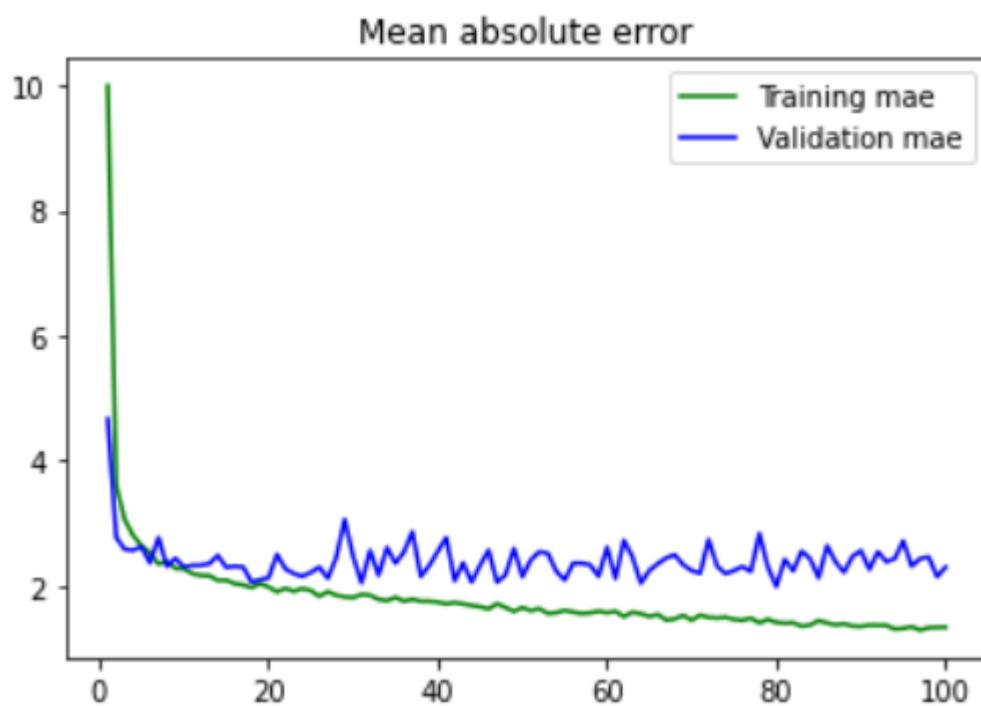


Рисунок 4

Последний тест был проведён с уменьшением количества блоков до 3-х.

Результат представлен на рисунке 5. Средняя ошибка оказалась равна 2.41,

выраженного минимума графика также не наблюдается, так что уменьшением блоков добиться снижением ошибки не удастся.

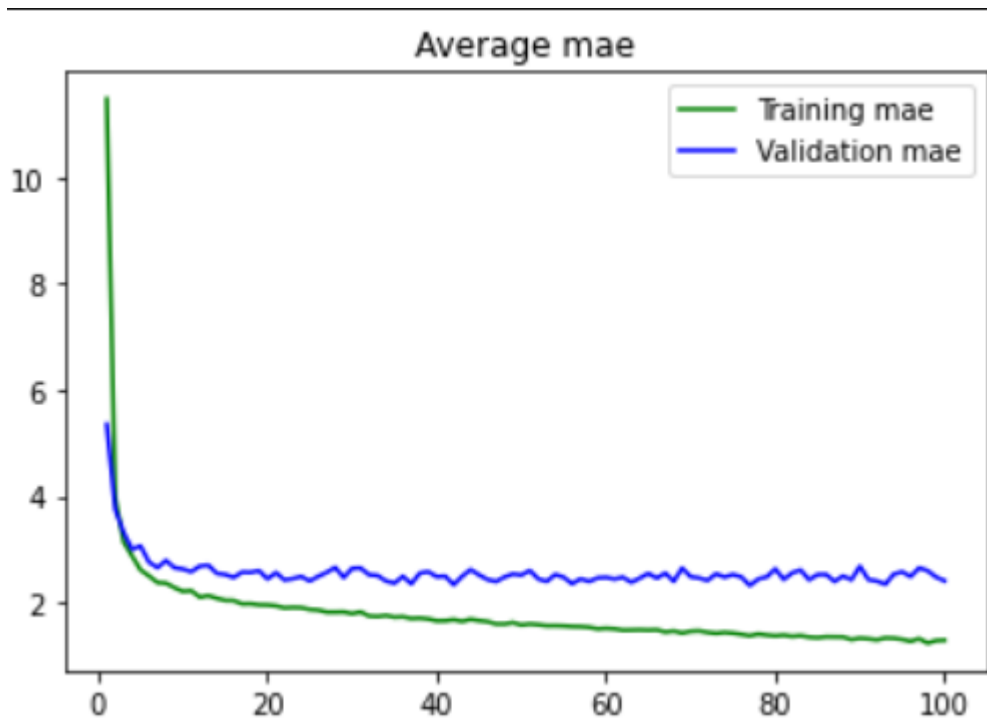


Рисунок 5

Выводы.

В ходе выполнения работы было реализовано предсказание медианной цены на дома в пригороде Бостона в середине 1970-х по разным данным. Был изучен метод перекрестной проверки по K блокам, а также влияние на него числа эпох и кол-ва блоков. Исследование было подкреплено результатами работы моделей, графиками ошибок моделей, а также усредненными графиками.