

**МИНОБРНАУКИ РОССИИ**  
**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ**  
**ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**  
**«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)**  
**Кафедра математического обеспечения и применения ЭВМ**

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

Студент гр. 8383

Преподаватель

\_\_\_\_\_

\_\_\_\_\_

Муковский Д.В.

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**

- Ознакомиться с задачей регрессии
- Изучить отличие задачи регрессии от задачи классификации
- Создать модель
- Настроить параметры обучения
- Обучить и оценить модели
- Ознакомиться с перекрестной проверкой

## **Требования**

- Объяснить различия задач классификации и регрессии
- Изучить влияние кол-ва эпох на результат обучения модели
- Выявить точку переобучения
- Применить перекрестную проверку по K блокам при различных K
- Построить графики ошибки и точности во время обучения для моделей, а также усредненные графики по всем моделям

## Ход работы

### Различия задач классификации и регрессии

- Процесс классификации моделирует функцию, с помощью которой данные прогнозируются в метках дискретных классов. С другой стороны, регрессия - это процесс создания модели, которая предсказывает непрерывное количество.
- Алгоритмы классификации включают дерево решений, логистическую регрессию и т. Д. Напротив, дерево регрессии и линейная регрессия являются примерами алгоритмов регрессии.
- Классификация предсказывает неупорядоченные данные, в то время как регрессия предсказывает упорядоченные данные.
- Регрессия может быть оценена с использованием среднеквадратичной ошибки. Наоборот, классификация оценивается путем измерения точности.

### Изучение влияния разных параметров на результат обучения модели

Для первого обучения были выбраны следующие параметры обучения: 100 эпох и перекрестной проверки 4 блока.

Были построены графики ошибки и средних абсолютных ошибок во время обучения для разных блоков, а также усредненные графики по всем блокам. Графики представлены на рис. 1-5.

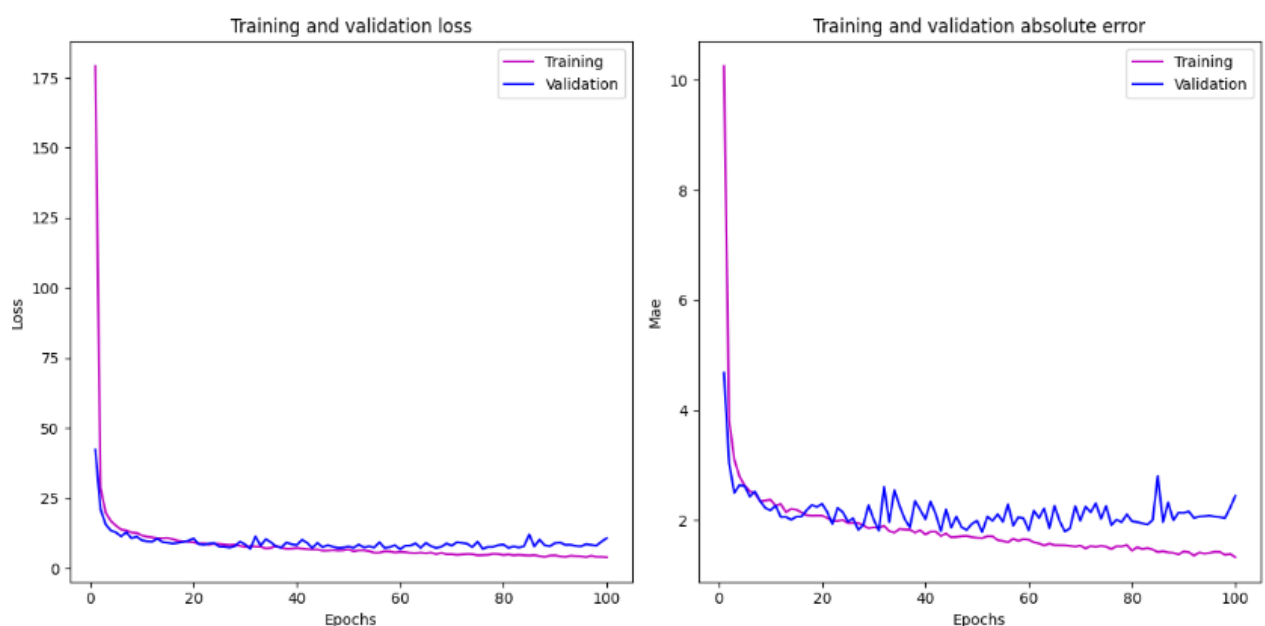


Рис. 1 – График оценки абсолютной ошибки для 1 блока

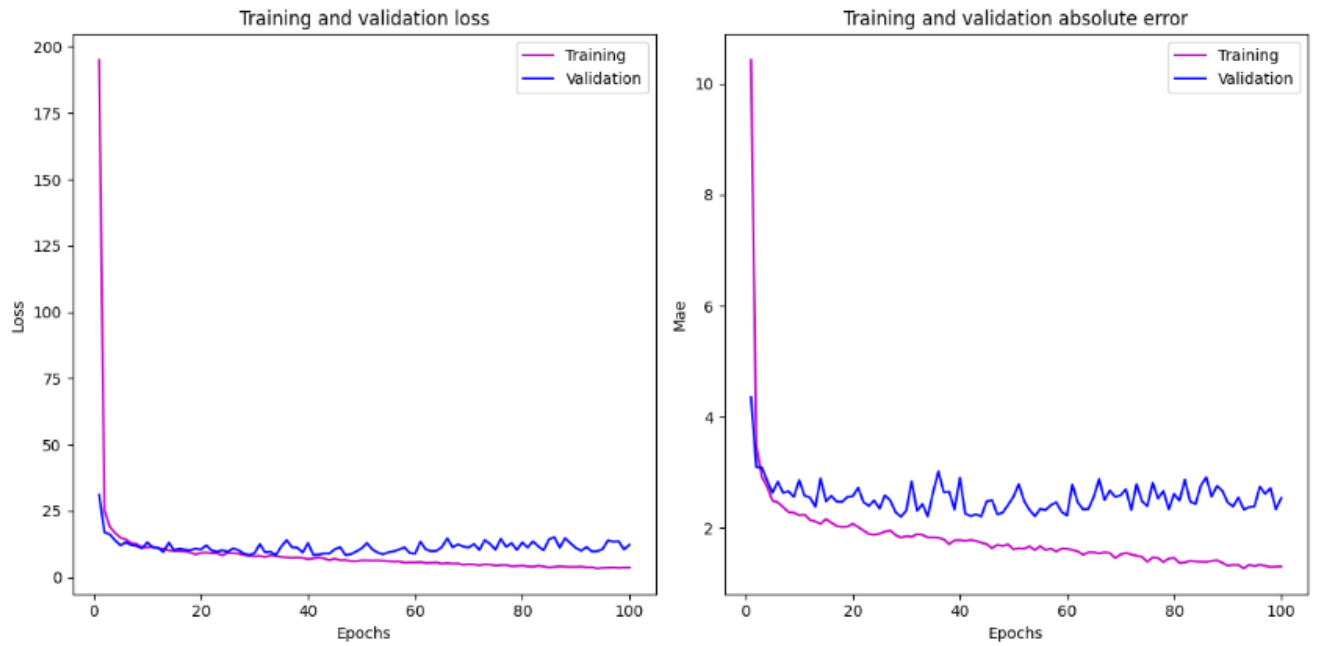


Рис. 2 – График оценки абсолютной ошибки для 2 блока

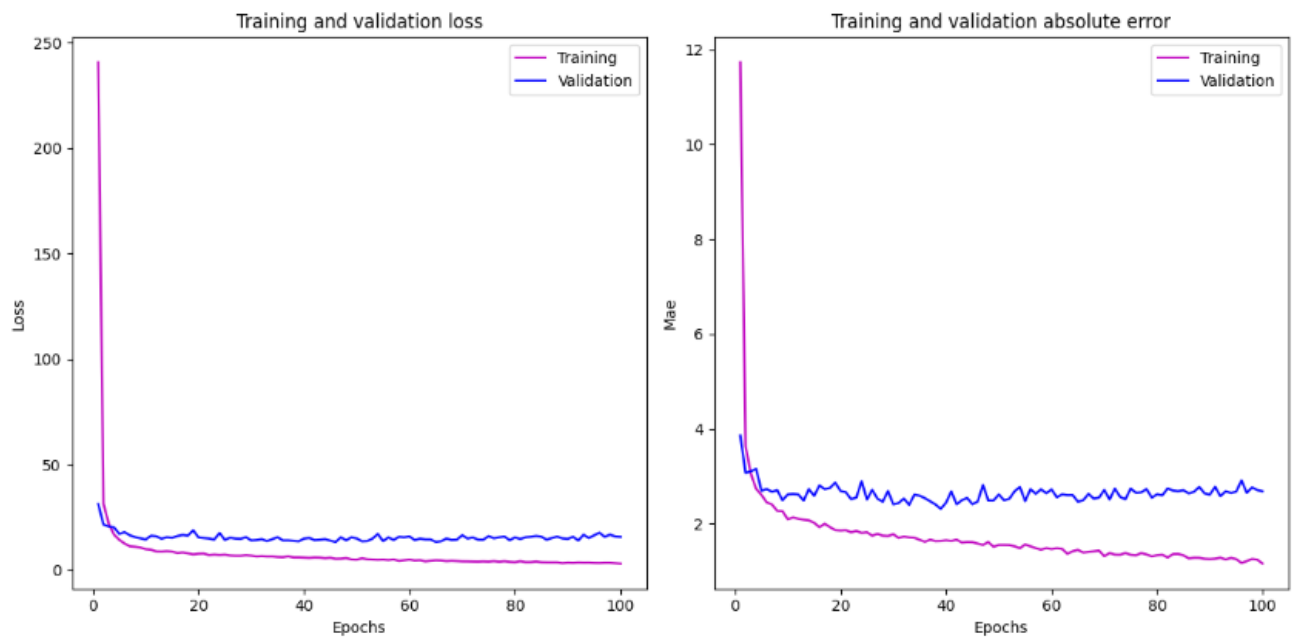


Рис. 3 – График оценки абсолютной ошибки для 3 блока

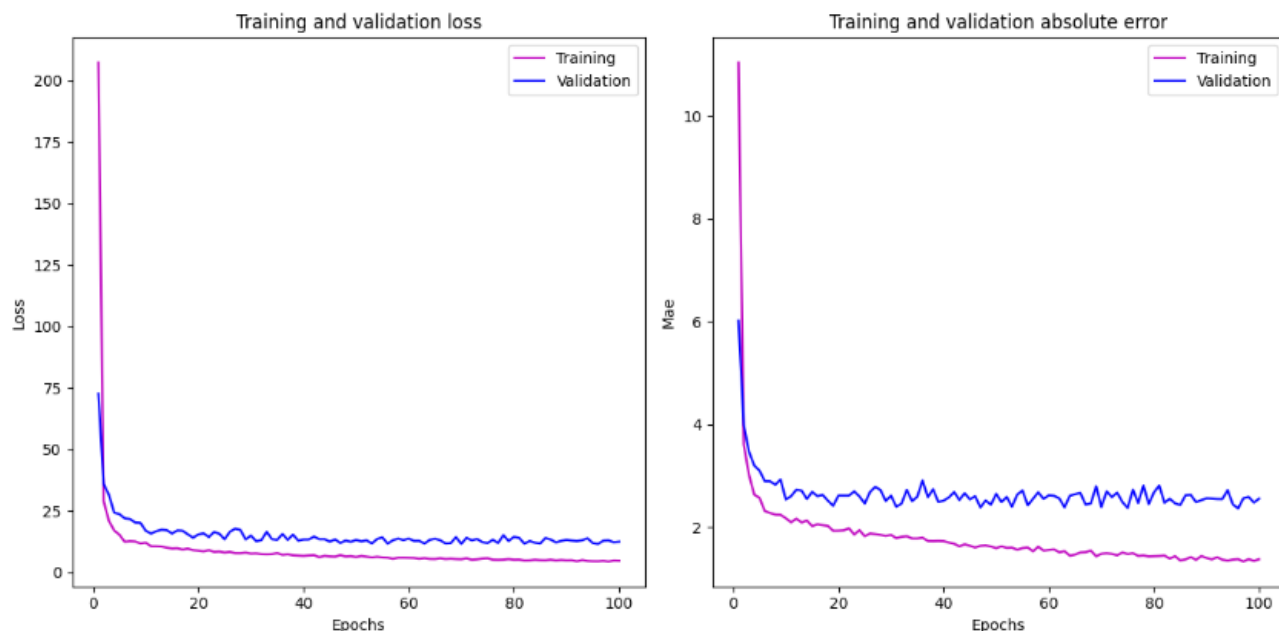


Рис. 4 – График оценки абсолютной ошибки для 4 блока

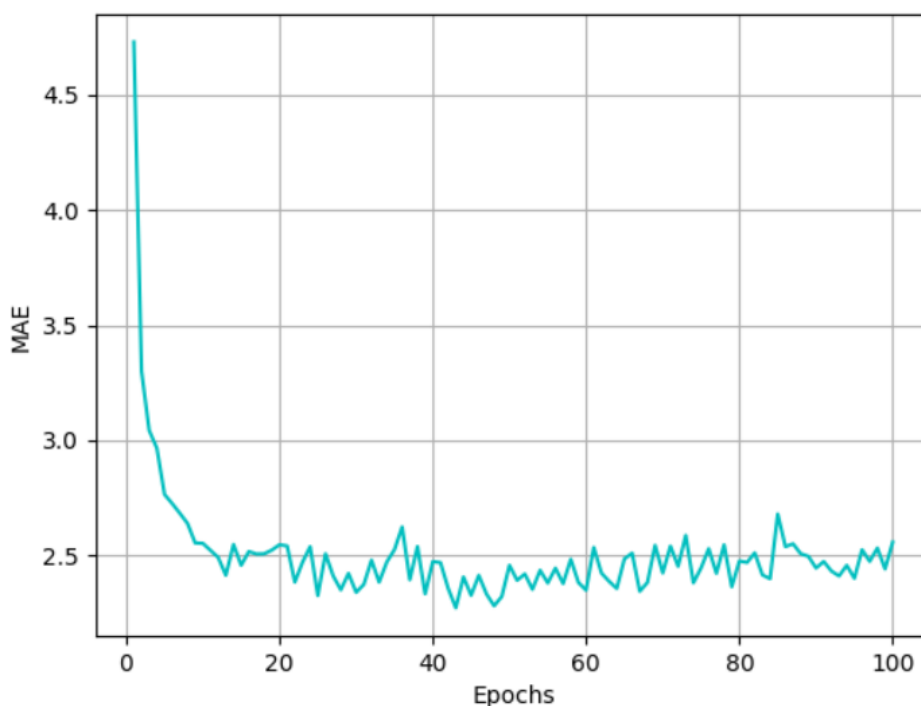


Рис. 5 – График среднего значения оценки средней абсолютной

Из последнего графика средней абсолютной ошибки можно увидеть, что на 50-й эпохе нейронная сеть переобучается, т.к. значение ошибки перестает уменьшаться. Такой же вывод можно сделать, глядя на графики потерь для четырех блоков.

Уменьшим количество эпох до 50. Графики представлены на рис. 6-9.

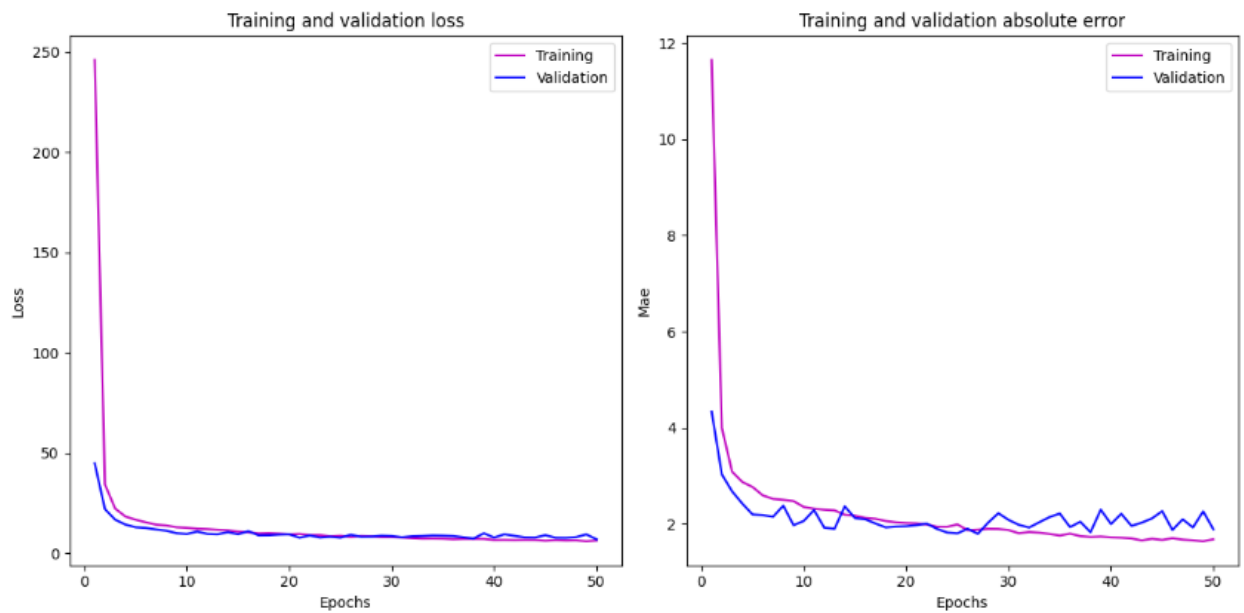


Рис. 6 – График оценки абсолютной ошибки для 1 блока

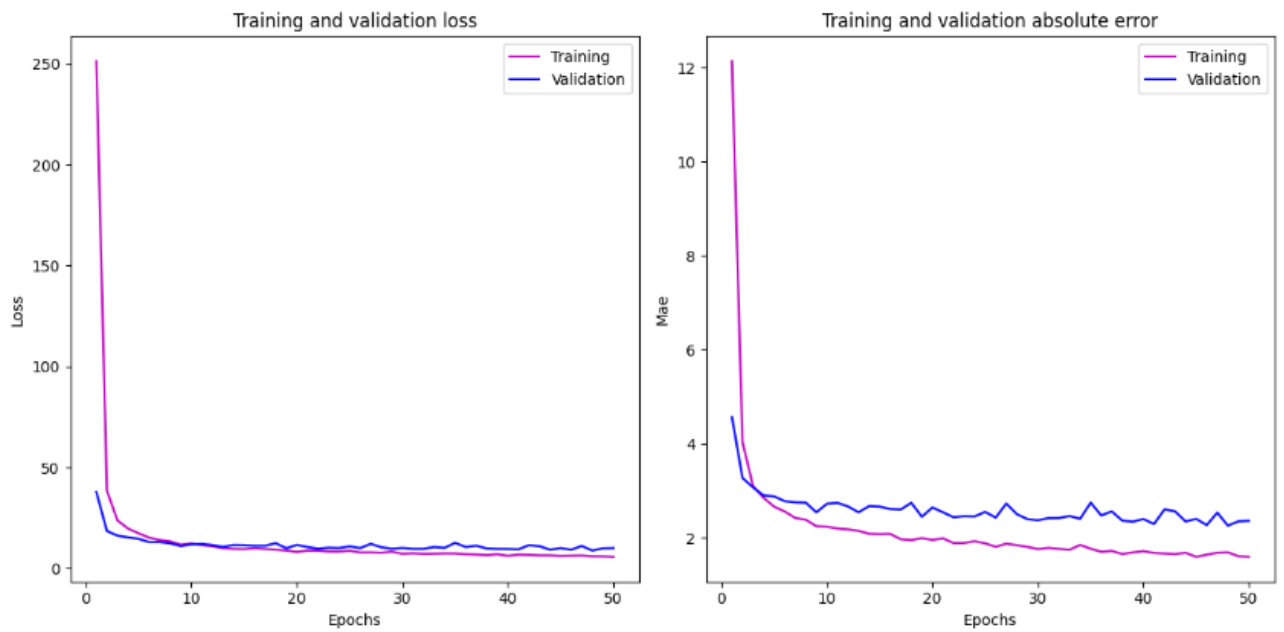


Рис. 7 – График оценки абсолютной ошибки для 2 блока

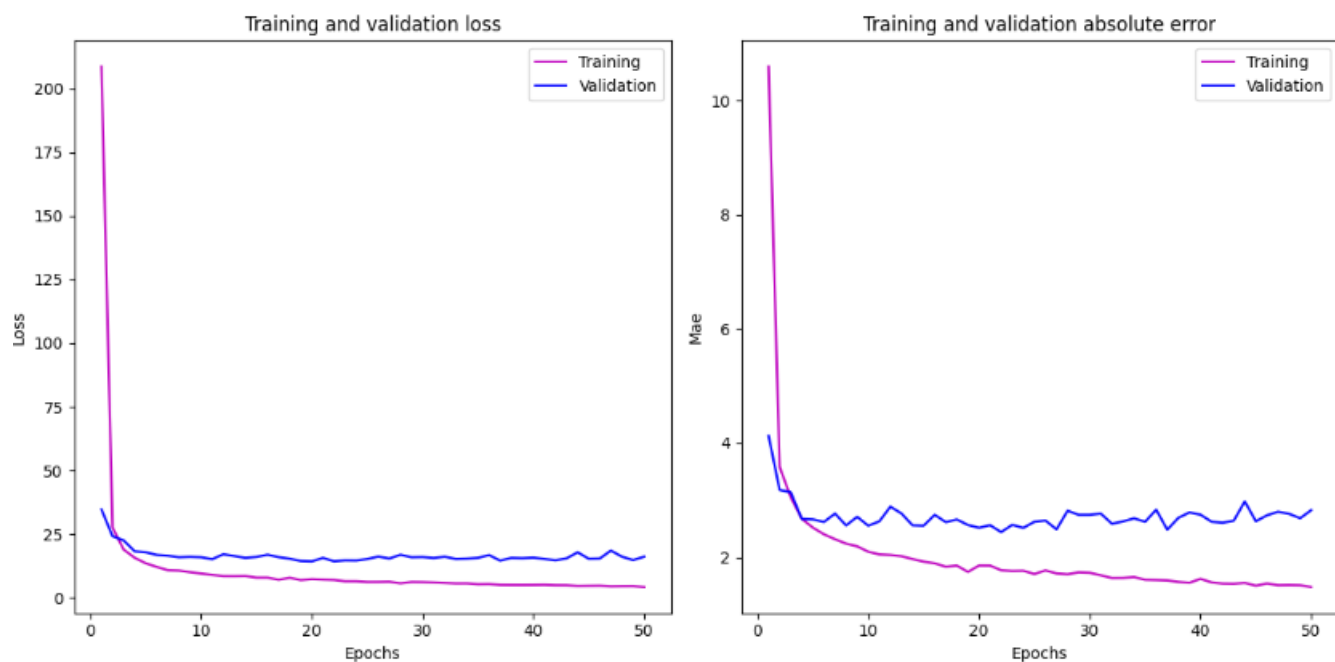


Рис. 8 – График оценки абсолютной ошибки для 3 блока

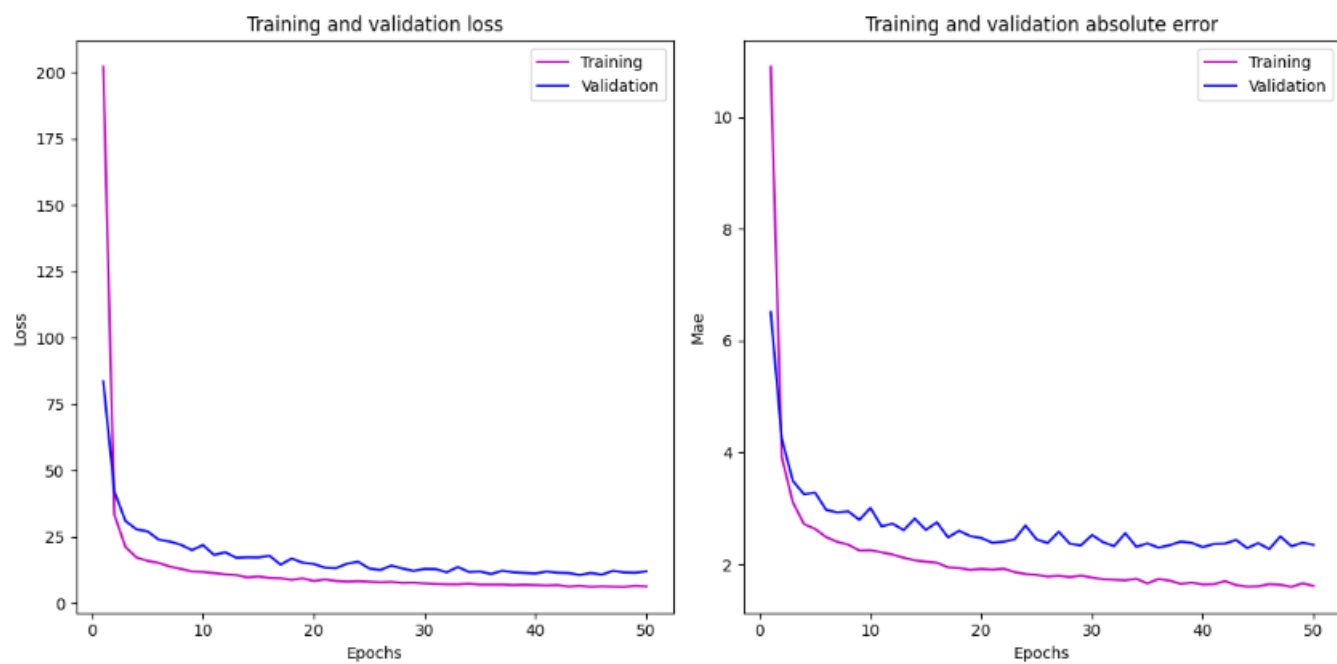


Рис. 9 – График оценки абсолютной ошибки для 4 блока

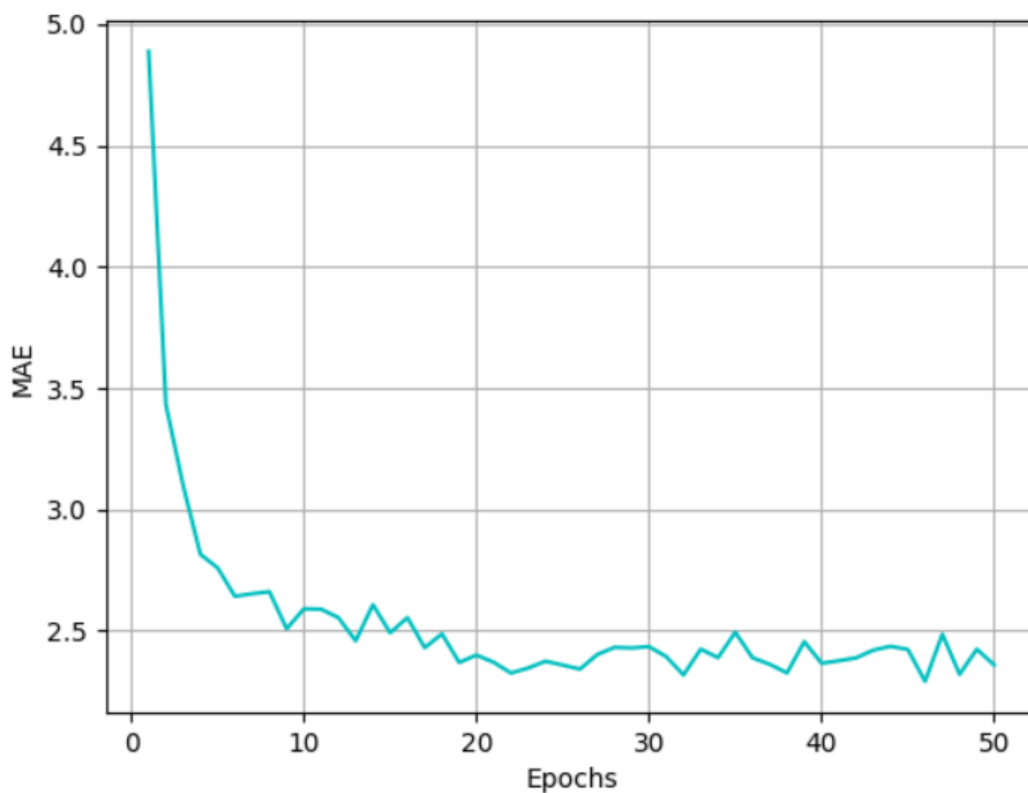


Рис. 10 – График среднего значения оценки средней абсолютной ошибки

По графику средней абсолютной ошибки видно, что теперь модель начала переобучаться на 20 эпохе. По графикам потерь также видно, что на валидационных данных ошибка после 20 эпохи не уменьшается.

Далее было увеличено количество блоков до 5. График среднего значения оценки средней абсолютной ошибки представлен на рис. 11.



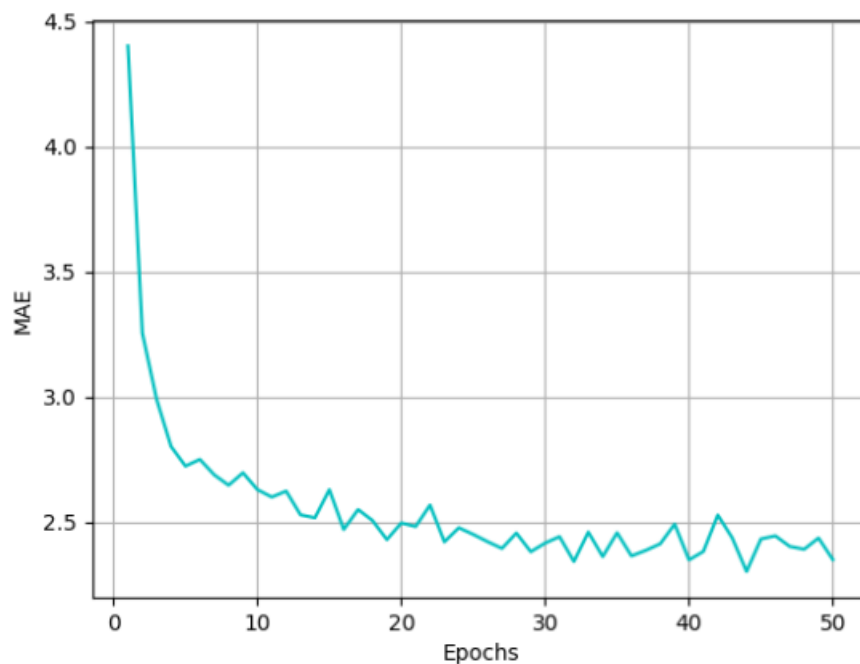


Рис. 11 – График среднего значения оценки средней абсолютной ошибки при  $k = 5$

По графику средней абсолютной ошибки видно, что теперь модель начала переобучаться на 40 эпохе, ошибка немного уменьшилась.

Далее было уменьшено количество блоков до 3. График среднего значения оценки средней абсолютной ошибки представлен на рис. 12.

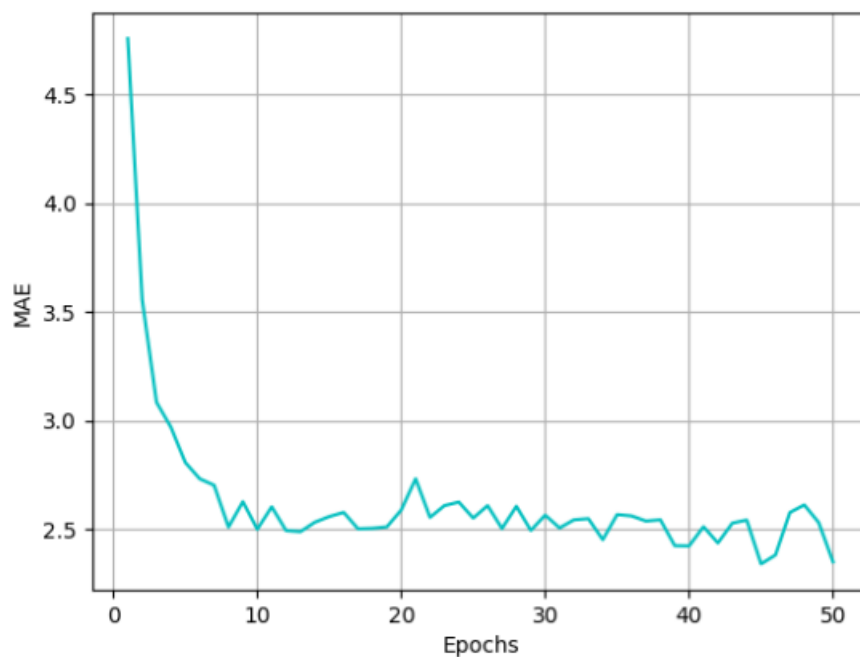


Рис. 12 – График среднего значения оценки средней абсолютной ошибки при  $k = 3$

Данная модель показала наихудший результат, ошибка самая большая. Полагаю, данные результаты вызваны тем, что тренировочных данных недостаточно.

Оптимальное количество блоков 5, а количество эпох обучения 20, при таких условиях ИНС показывает наименьшую ошибку.

## **Выводы**

В ходе выполнения лабораторной работы была изучена реализация регрессии в машинном обучении для решения задачи предсказания медианной цены на дома по различным показателям. Также было изучено влияние числа эпох и количества блоков для перекрестной проверки на результат обучения нейронной сети.