



UNIVERSIDADE ESTADUAL DE MARINGÁ
DEPARTAMENTO DE ENGENHARIA QUÍMICA
CURSO DE ENGENHARIA ELÉTRICA

SISTEMAS INTELIGENTES - 6694

Avaliação 2

Professor:

Rafael Krummenauer

Alunos:

1. Ana Beatriz Louzada Piassa – RA: 101216
2. Caio de Andrade Caetano – RA: 98425
4. Hugo Fusinato – RA: 101201
3. Lucas Barzotto – RA: 98893

Este relatório tem como objetivo apresentar o treinamento de uma rede neural a partir de atributos e do conjunto de dados Boston Housing. A comparação se deu da seguinte forma: comparamos o desempenho de duas redes neurais treinadas, uma rede com todos os atributos e outra cujos atributos foram retirados progressivamente para fins de comparação. No final, após fazer alguns testes com algumas combinações de parâmetros, selecionamos os que dentre eles apresentaram maior predição dos dados

Os dados estudados foram:

1. **CRIM**: Taxa de criminalidade per capita.
2. ZN: Proporção de terrenos residenciais zoneados para lotes com mais de 25,000 pés quadrados.
3. INDUS: Proporção de acres de negócios não-varejistas por cidade.
4. CHAS: Variável dummy Charles River(= 1 se o trato limita rios; 0 caso ocorra contrário).
5. NOX: Concentração de óxidos nítricos (partes por 10 milhões).
6. **RM**: Média de número de quartos por habitação.
7. AGE: Proporção de unidades ocupadas pelos donos construídas antes de 1940.
8. DIS: Distâncias ponderadas à cinco centros de emprego de Boston.
9. RAD: Índice de acessibilidade a rodovias radiais.
10. TAX: Taxa de imposto de propriedade de valor total por \$10,000.
11. PTRATIO: Proporção aluno-professor por cidade.
12. B: $1000 * (B_k - 0.63) ** 2$ onde B_k é a proporção de pessoas negras por cidade.
13. **LSTAT**: Percentual de status [inferior da população](#).

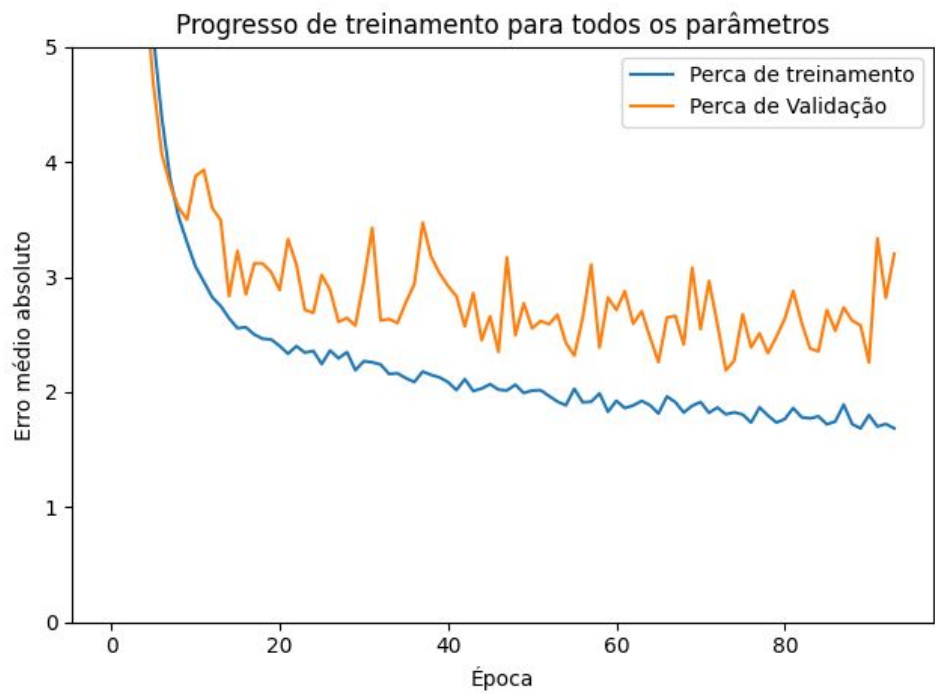
Após alguns testes com várias combinações diferentes de parâmetros vimos que trabalhar com os dados de criminalidade (CRIM), quartos por habitação (RM) e percentual de status da população (LSTAT). Conseguindo obter erros menores treinando apenas com esses parâmetros do que com todos em alguns casos de treinamento.

Foram usadas 4 camadas, sendo uma inicial de 64 neurônios, a final com uma saída e duas hidden layers com 64 neurônios com ativação do tipo ReLU (Função de ativação linear retificada) para o treinamento.

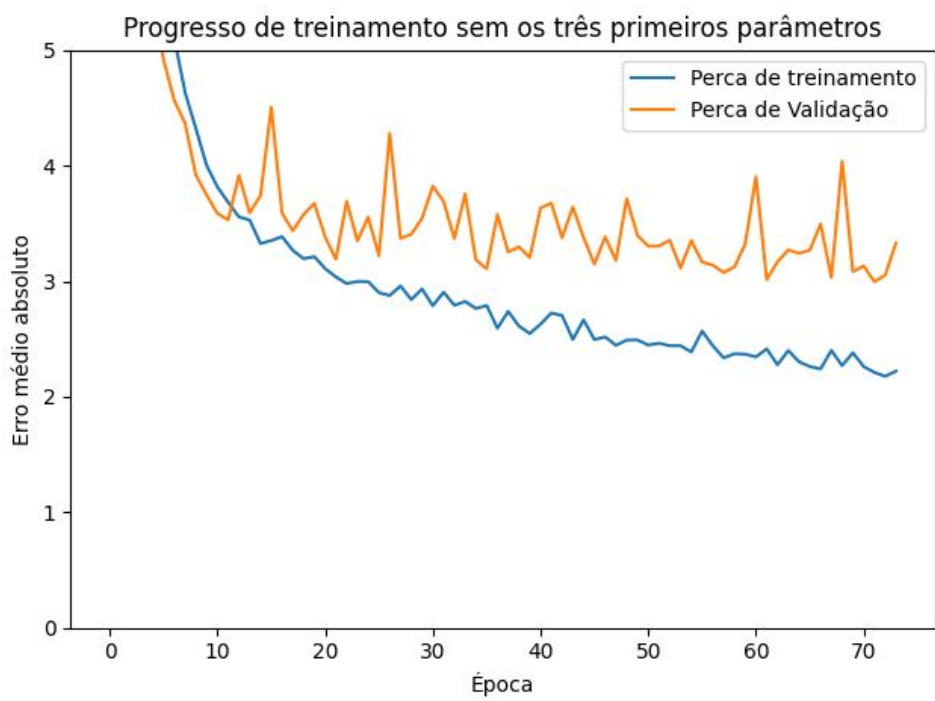
Segue abaixo os resultados em termos de evolução do erro de treinamento e validação para cada caso em específico.

1. Gráficos do Progresso de Treinamento

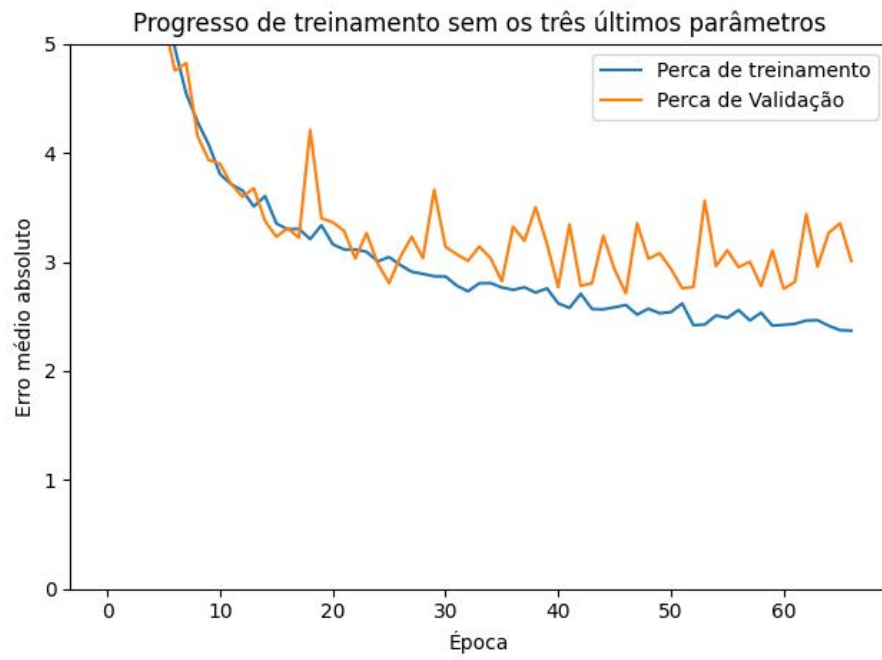
a.



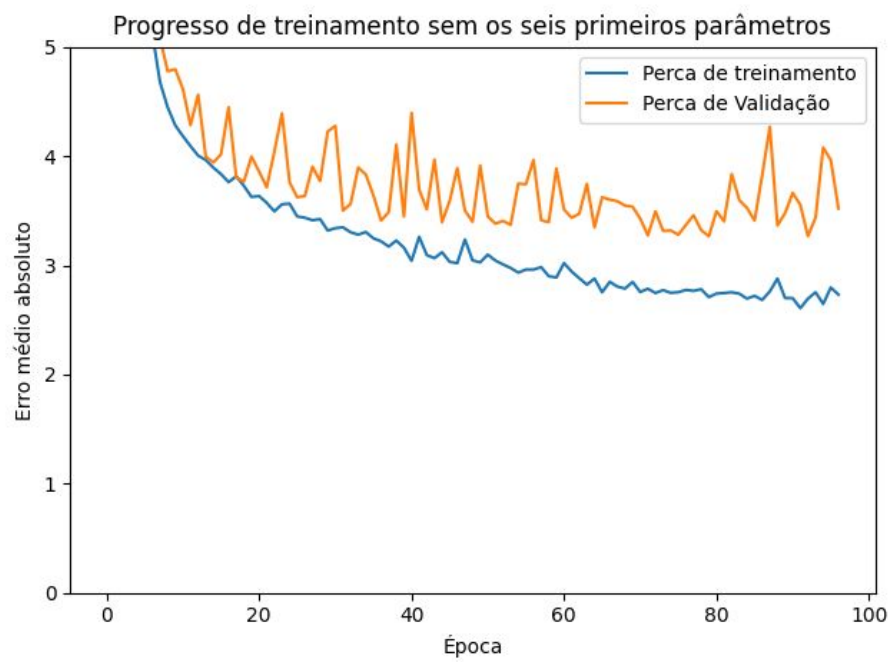
b.



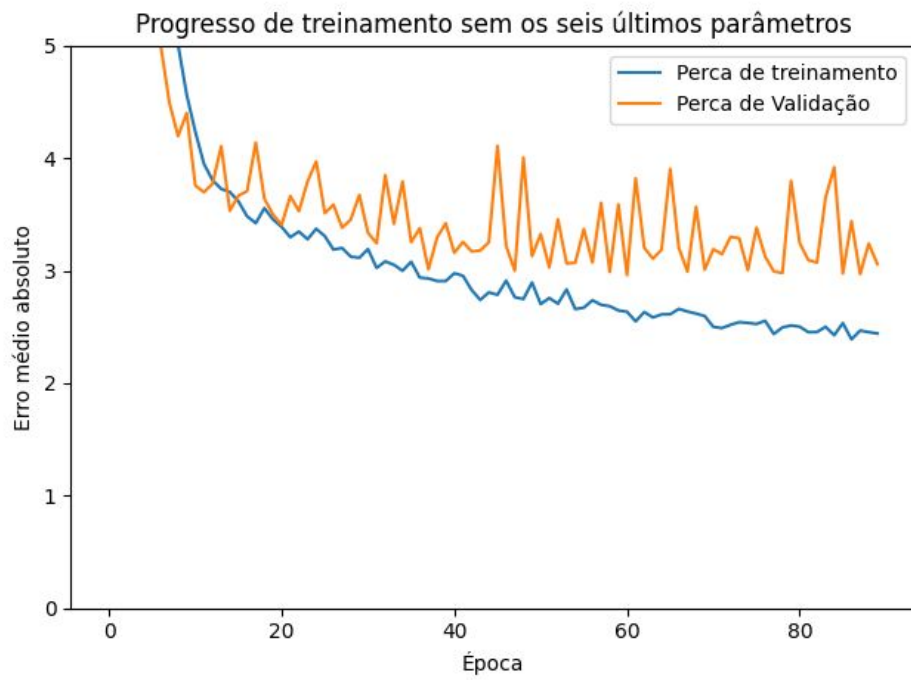
c.



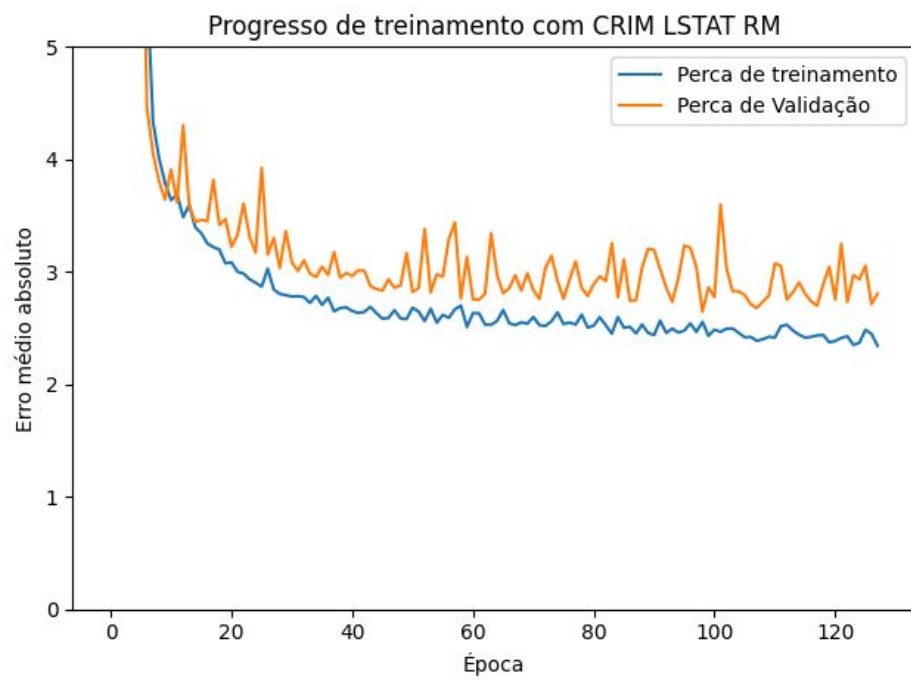
d.



e.

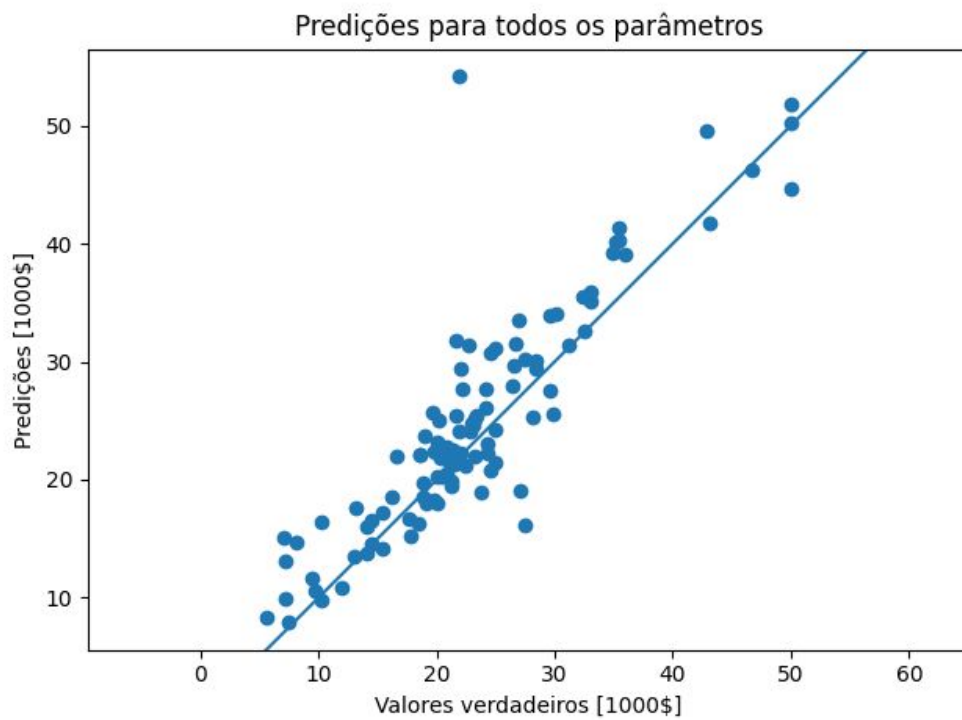


f.

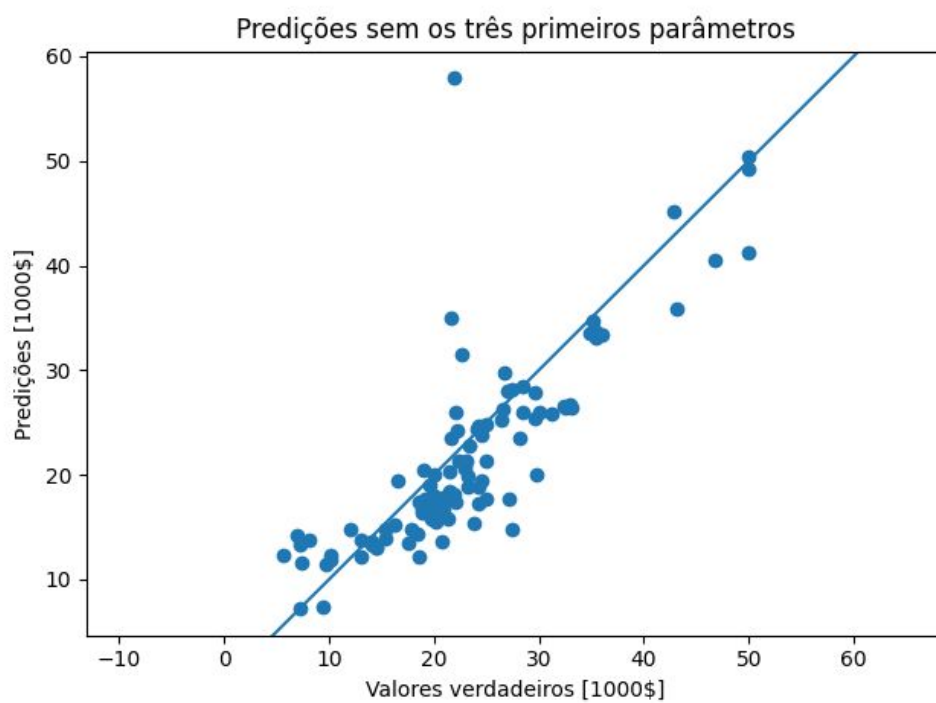


2. Predição do Progresso de Treinamento

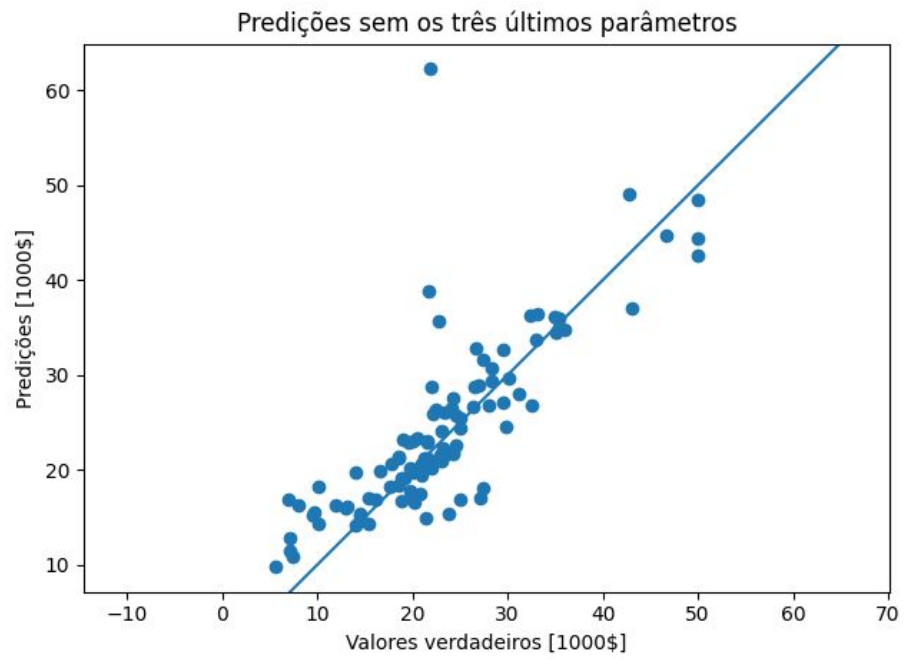
a.



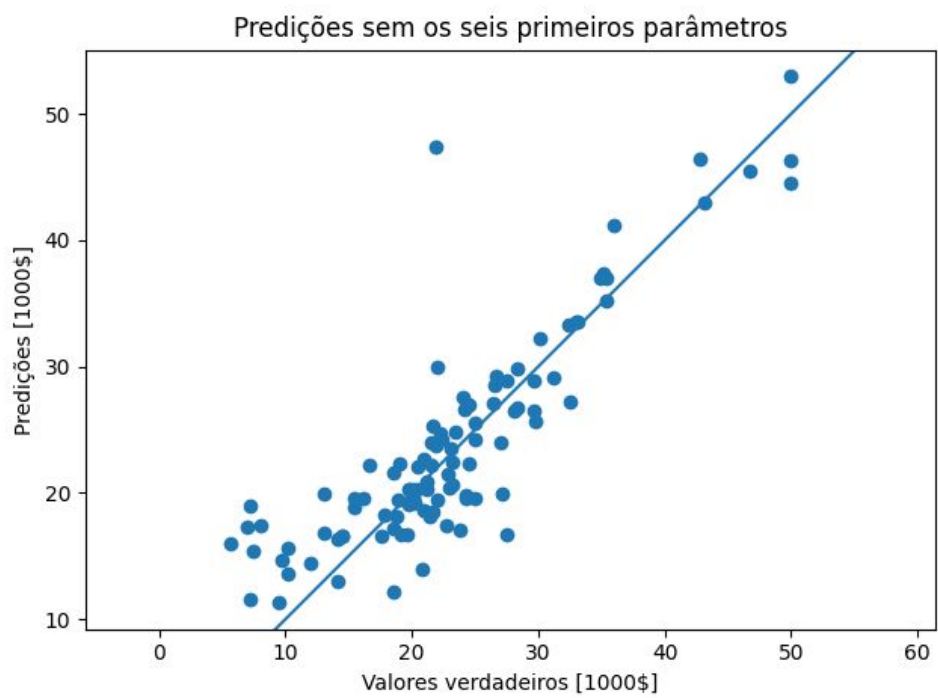
b.



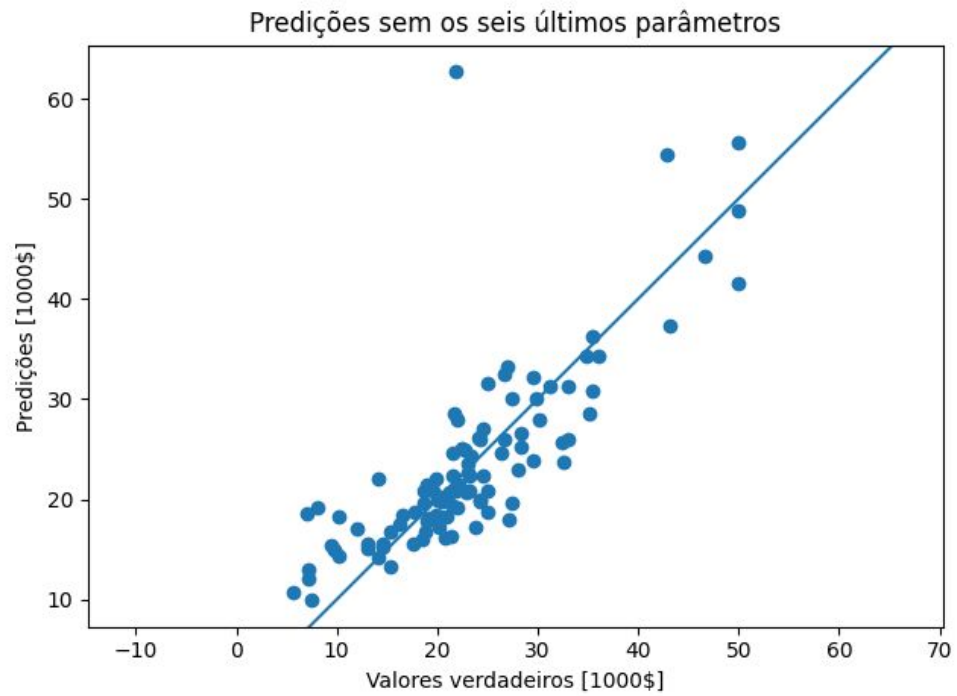
c.



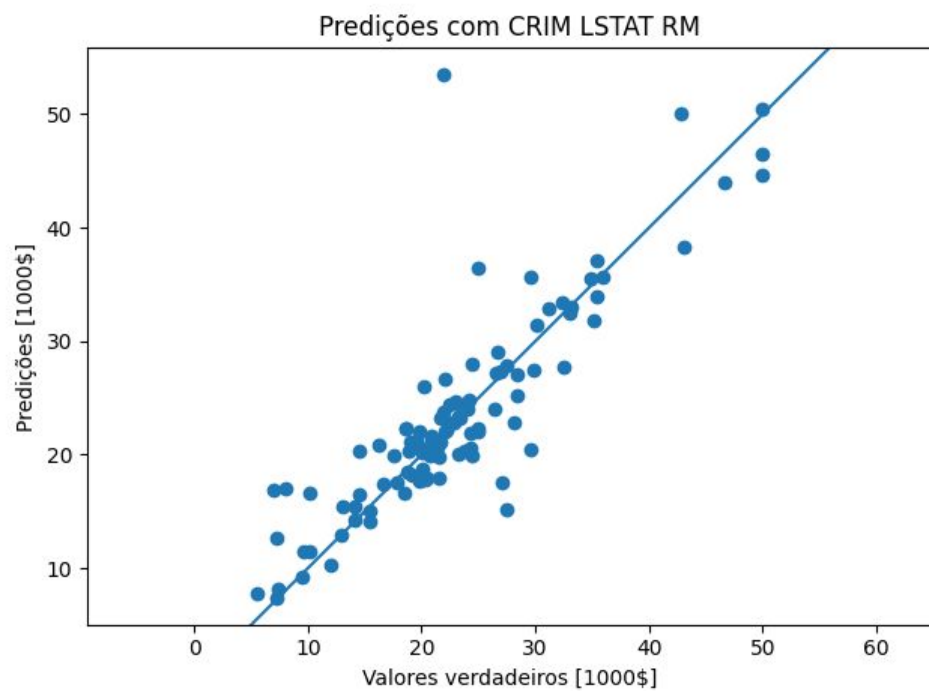
d.



e.

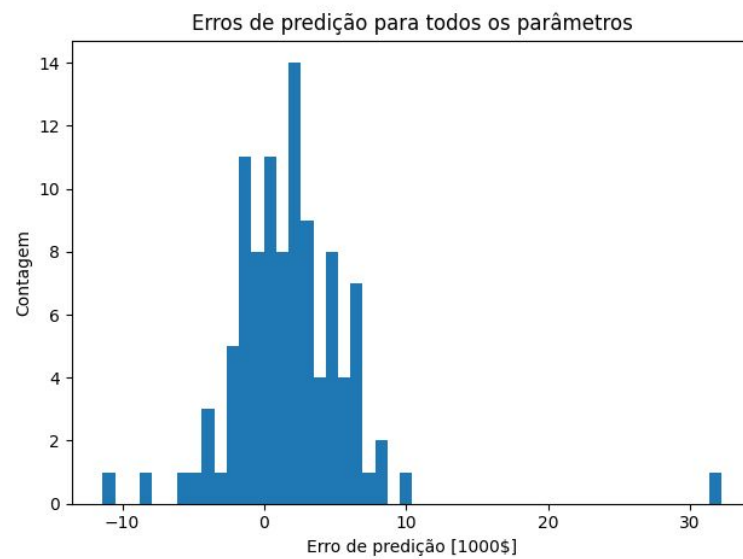


f.

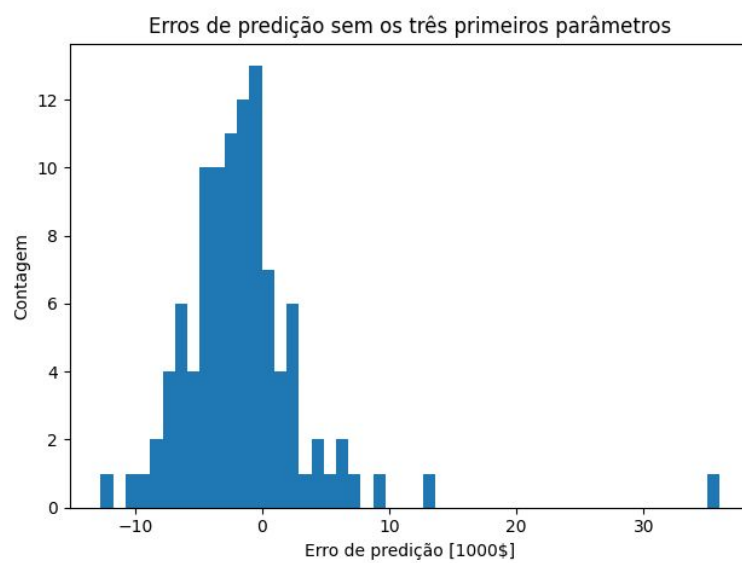


3. Erros do Progresso de Treinamento

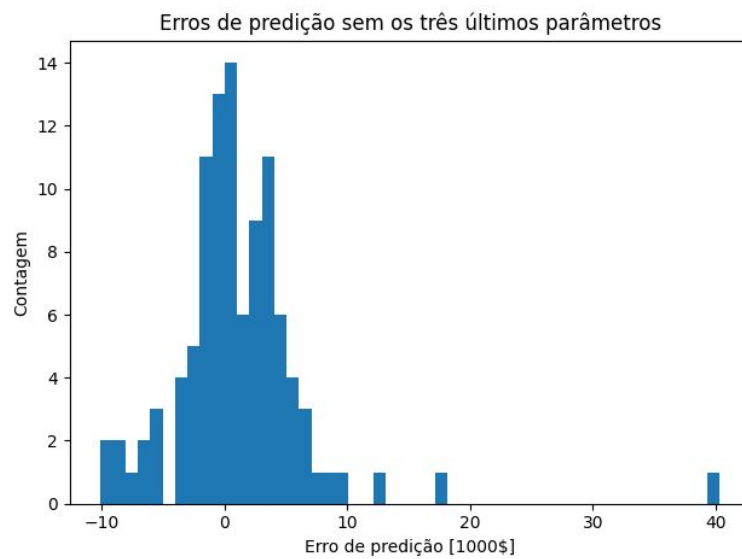
a.



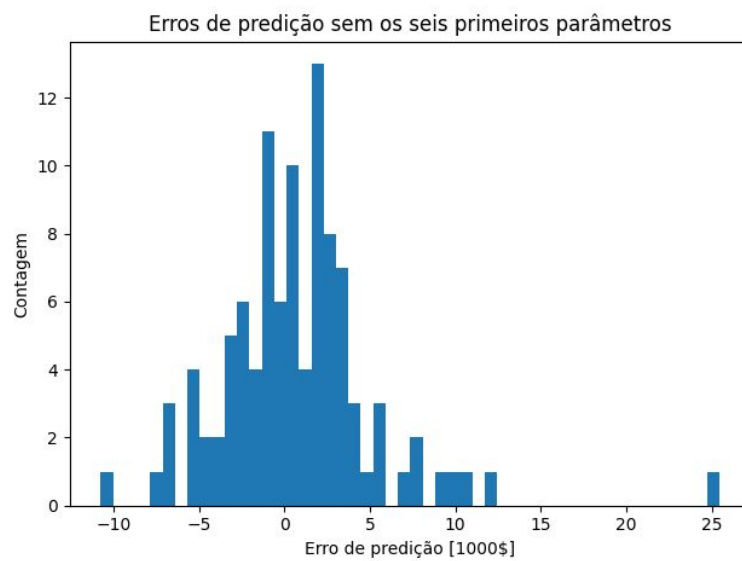
b.



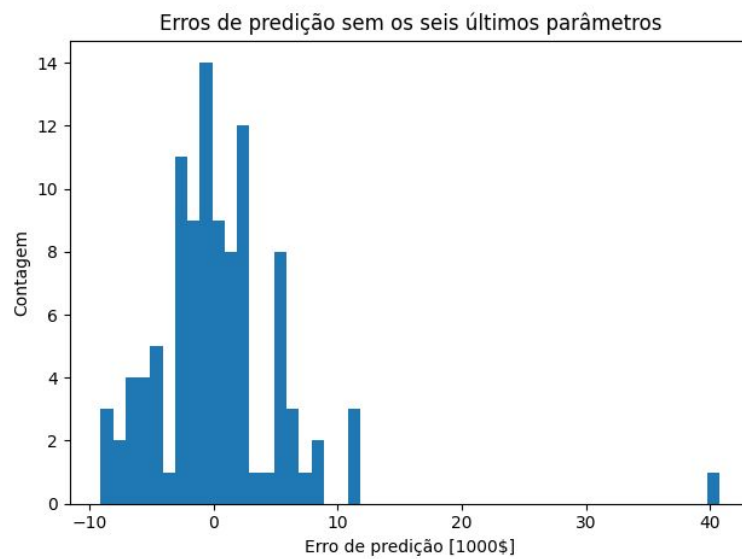
c.



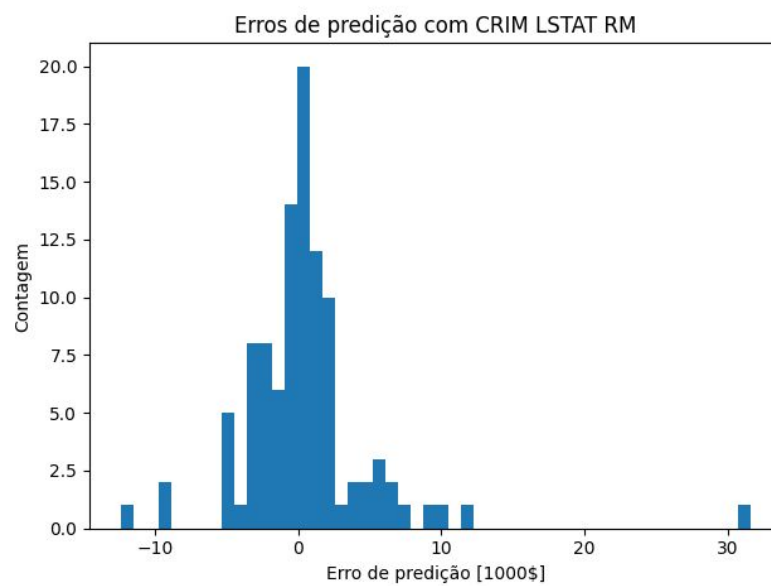
d.



e.

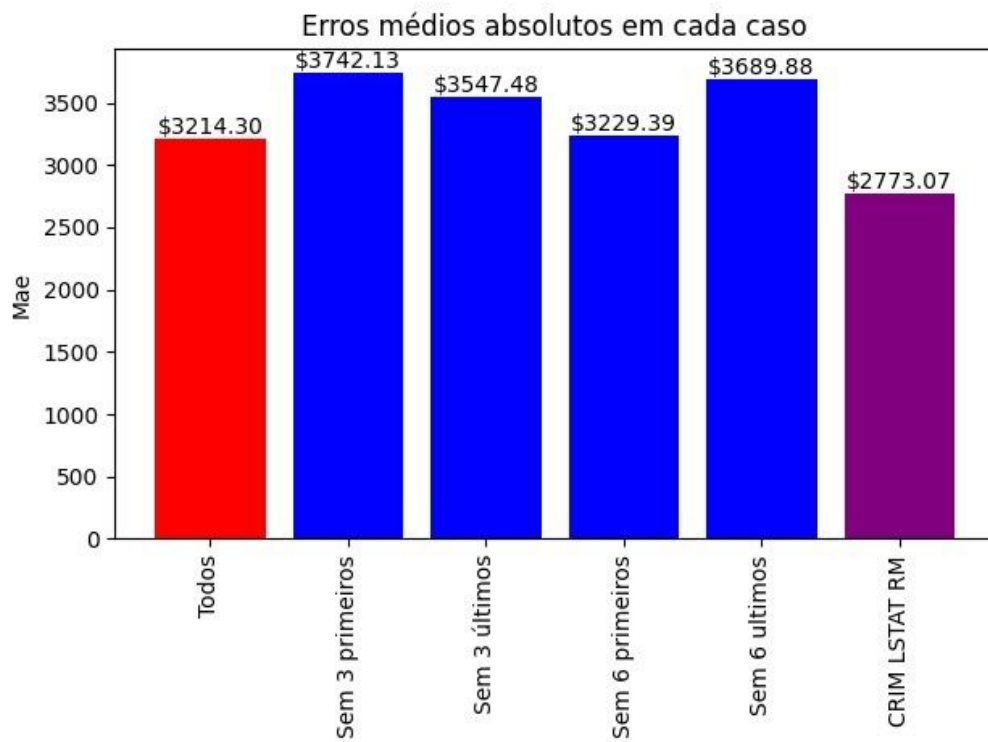


f.

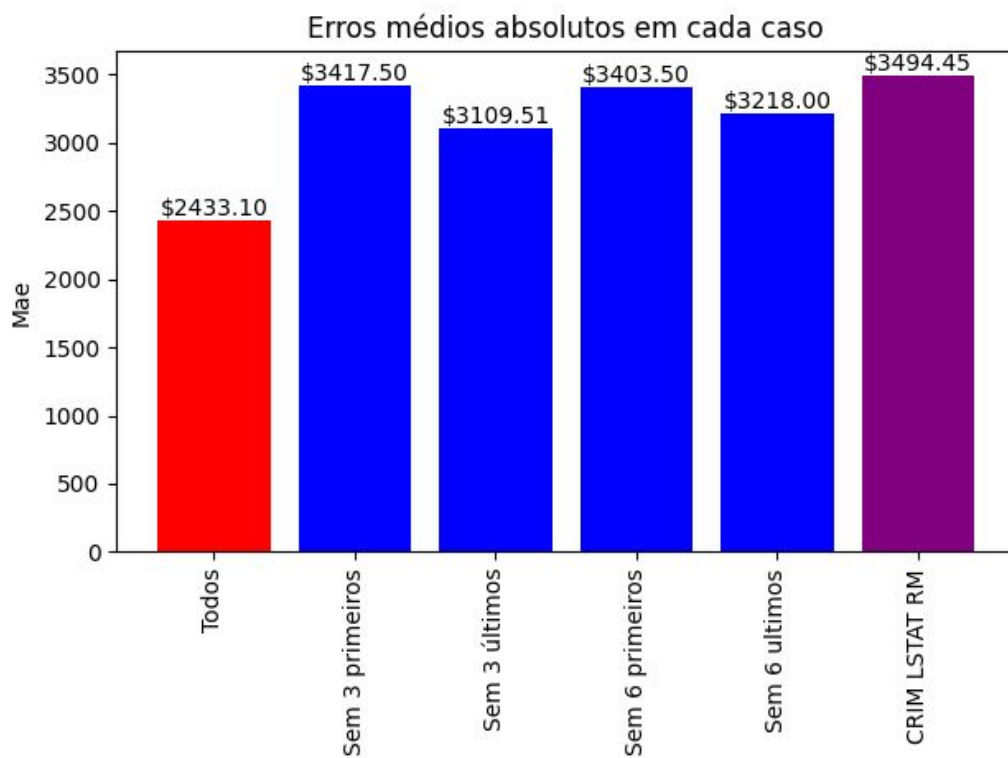


4. Comparação dos erros de alguns treinamentos

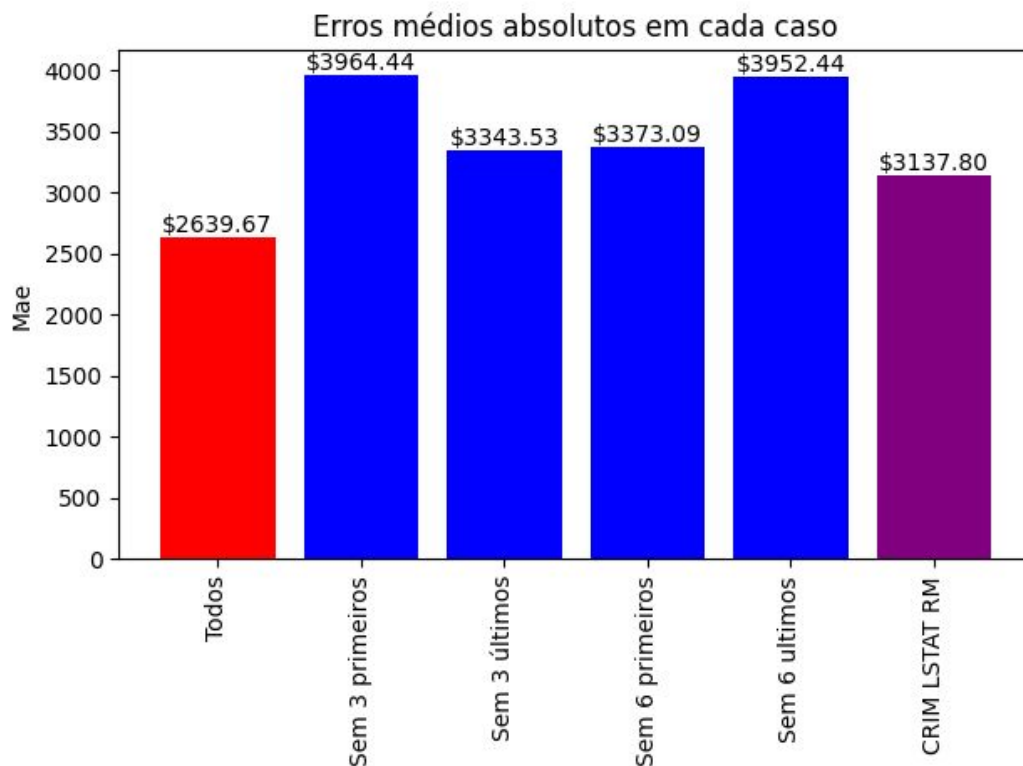
a.



b.



c.



Após simular alguns casos de teste que alteraram os dados diferentemente na função random, vimos que no geral quanto mais parâmetros temos de treinamento geralmente melhor é nossa predição e menor o erro.

Como a randomização muda o sistema e talvez possa gerar um overtraining (apesar de existir uma função de early_stop para o treinamento) em cada iteração obtivemos uma correlação diferente entre o uso de dados e nem sempre tivemos um padrão que mostrasse quais parâmetros podem ser mais efetivos em treinar a rede neural.

Pode-se concluir que a própria rede consegue identificar quais dos parâmetros são menos relevantes para prever o preço das casas e colocar uma menor importância matemática para a variação desses valores. Quanto mais parâmetros confiáveis sejam inseridos em uma rede neural, maior é a probabilidade dela prever melhor os dados.