

Desafio 01 - REDES NEURAIS ARTIFICIAIS (APRENDIZADO PROFUNDO)

Hanna Tatsuta Galassi

2025-09-17

Table of contents

1	Introdução	1
2	Discussão	1
2.1	Topologia mínima	1
2.2	Utilização de diferentes funções de ativação	2
3	Conclusões	4

1 Introdução

Este relatório apresenta os resultados do desafio de modelagem de uma função univariada desconhecida usando uma Rede Neural Artificial do tipo Perceptron de Múltiplas Camadas (MLP). O objetivo principal do estudo é encontrar a topologia mínima (o número de camadas e neurônios) necessária para modelar a relação entre a feature `xgrd` e a variável alvo `ymean`.

Além de determinar a arquitetura ideal, o estudo investiga a influência de diferentes funções de ativação—ReLU, tangente hiperbólica (`tanh`) e sigmoide—no desempenho do modelo. Para garantir a robustez das conclusões, cada configuração testada foi executada múltiplas vezes para avaliar a variabilidade dos resultados. O relatório discute as principais descobertas obtidas a partir dessas análises, apresentando uma discussão sobre a eficiência e a capacidade de generalização de cada modelo.

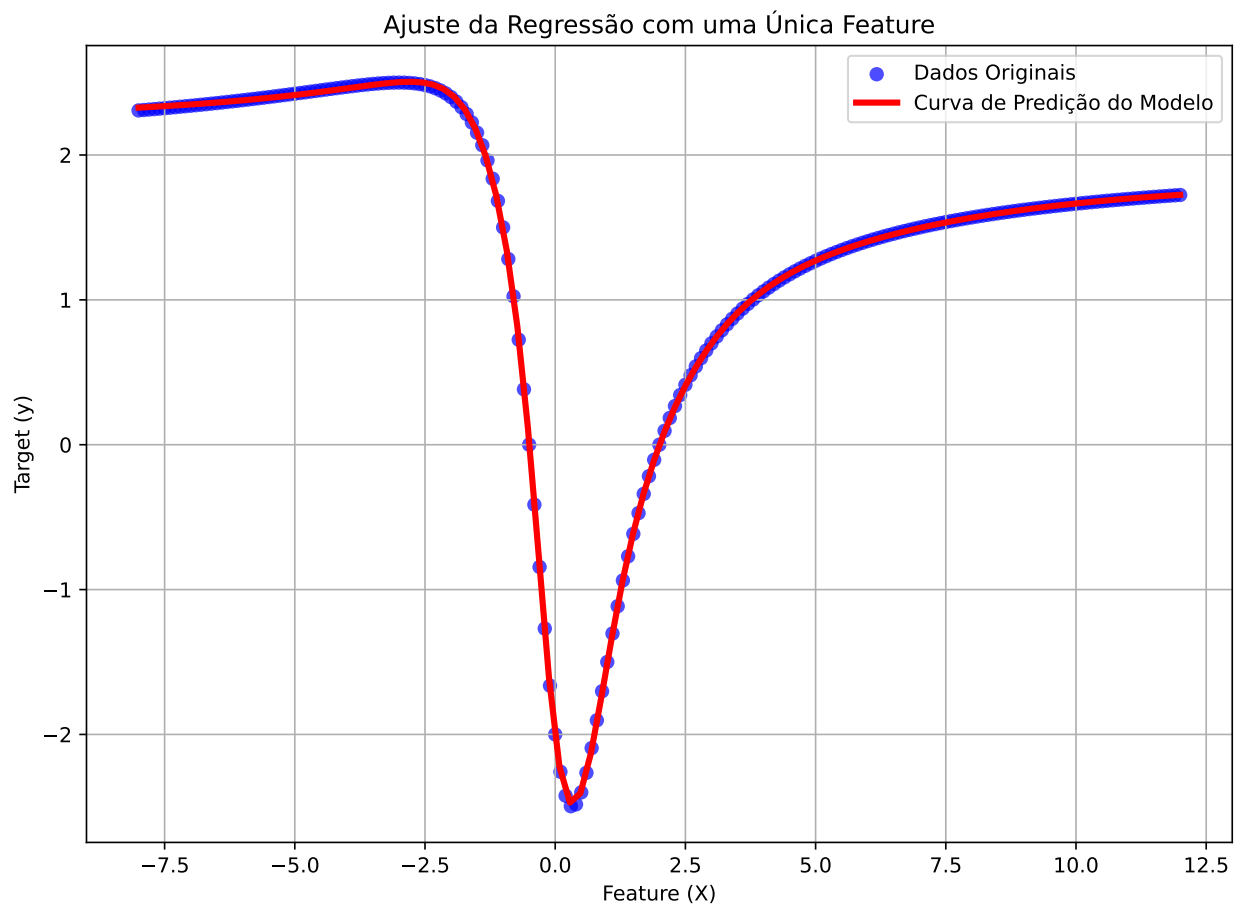
2 Discussão

O processo de desenvolvimento de um modelo de aprendizado profundo envolve a escolha de hiperparâmetros que influenciam diretamente seu desempenho. Esta seção detalha a metodologia utilizada para encontrar a topologia ideal da rede neural, bem como a avaliação do impacto de diferentes funções de ativação e da variabilidade inerente ao treinamento.

2.1 Topologia mínima

A busca pela topologia mínima foi realizada por meio de uma abordagem iterativa. Após diversos testes, a topologia de uma única camada oculta com 8 neurônios e função de ativação sigmoide foi identificada como a arquitetura mais simples e eficiente para modelar a relação entre `xgrd` e `ymean` ao longo de 2000 passos. Essa arquitetura demonstrou ser suficiente para capturar a forma da função de maneira satisfatória, sem a necessidade de camadas adicionais ou maior complexidade.

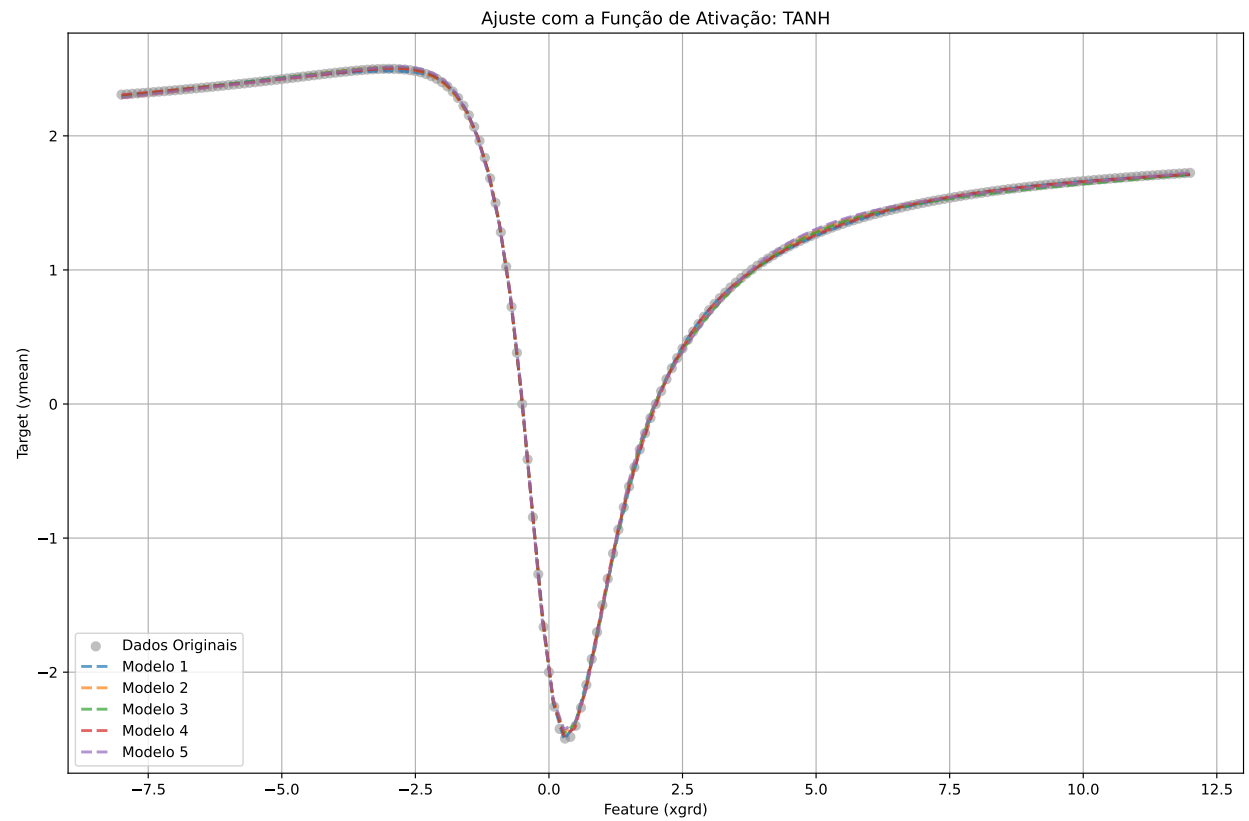
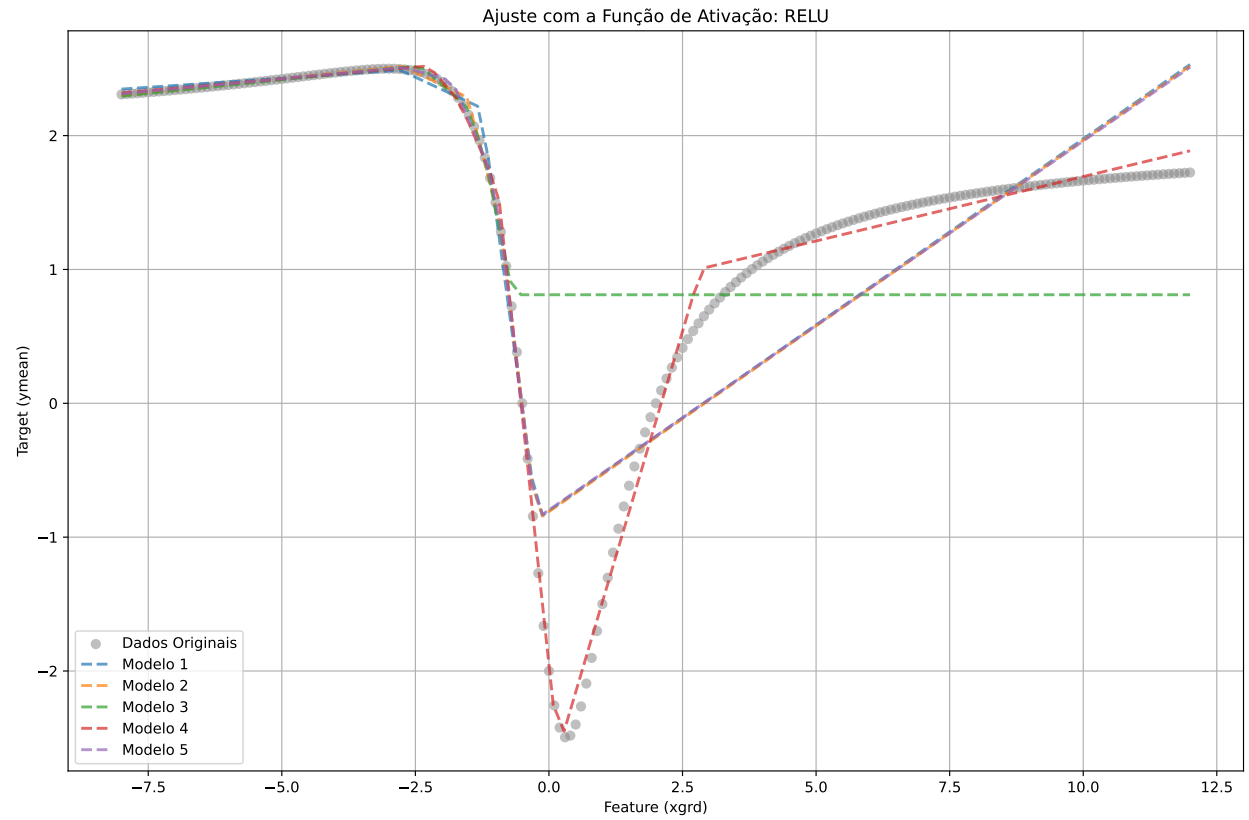
A seguir, é apresentado um gráfico que ilustra o ajuste da topologia mínima encontrada (linha vermelha) em relação à curva entre `xgrd` e `ymean`.

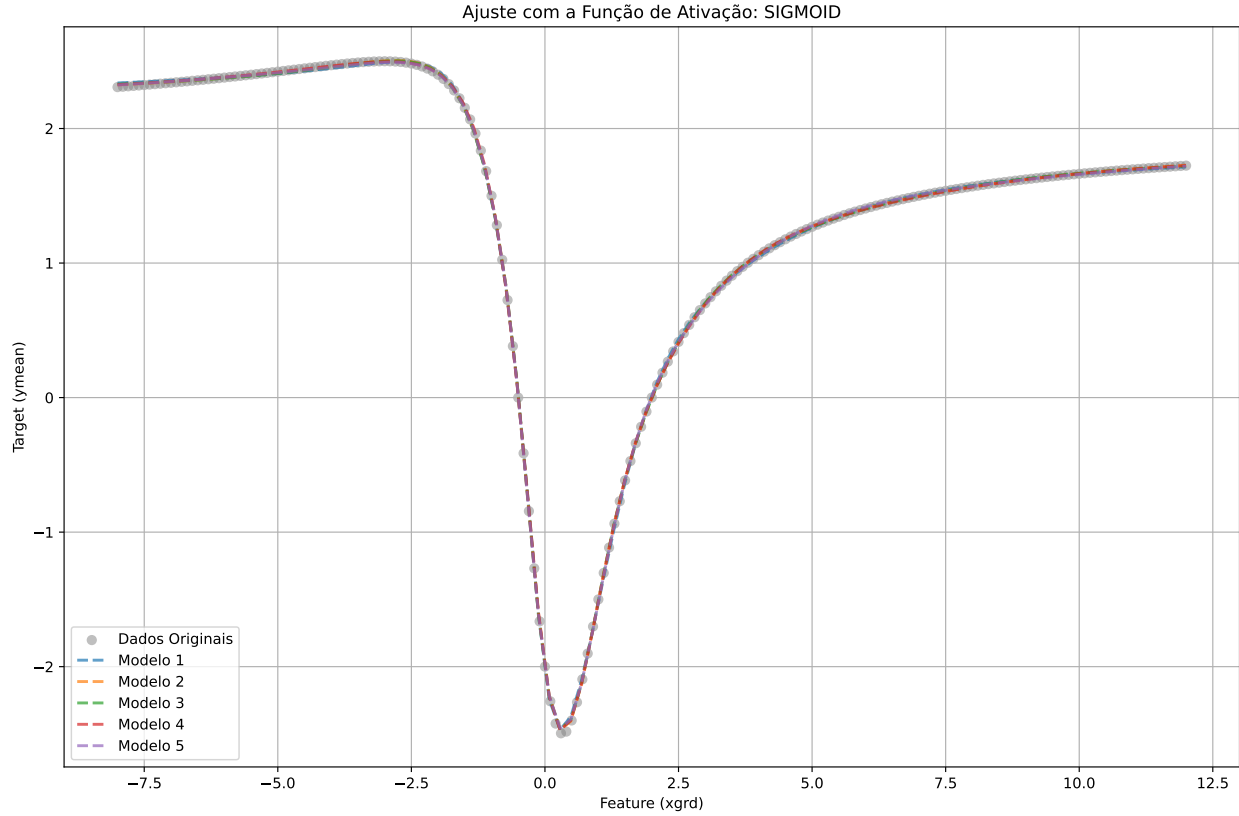


2.2 Utilização de diferentes funções de ativação

Com a topologia mínima definida, a análise foi expandida para avaliar a influência de três funções de ativação (ReLU, tangente hiperbólica (tanh) e sigmoide) no desempenho do modelo. Para cada função, foram treinados cinco modelos com as mesmas condições iniciais para analisar a variabilidade dos resultados. Os gráficos a seguir ilustram o ajuste dos cinco modelos treinados com cada função de ativação.

```
WARNING:tensorflow:5 out of the last 9 calls to <function TensorFlowTrainer.make_predict_function.<local>
WARNING:tensorflow:6 out of the last 12 calls to <function TensorFlowTrainer.make_predict_function.<local>
```





A topologia de 8 neurônios com a função de ativação ReLU demonstrou um comportamento consistente, mas resultou em um ajuste insatisfatório. Embora todos os cinco modelos tenham convergido para a mesma solução, a natureza linear da função ReLU para valores positivos de entrada impediu que a rede neural capturasse a forma completa dos dados originais. O resultado foi um modelo que se ajustou bem apenas à primeira parte da curva, falhando em modelar a segunda metade.

Em contraste, as funções sigmoide e tangente hiperbólica (tanh) demonstraram um desempenho superior e, principalmente, uma consistência notável. Em ambos os casos, todos os cinco modelos treinados convergiram para a mesma solução, ajustando-se de forma precisa e satisfatória aos dados originais. Essa consistência, evidenciada pela sobreposição das curvas, comprova que a topologia de 8 neurônios, combinada com as funções sigmoide ou tanh, é robusta e capaz de encontrar a solução ideal de forma confiável para este problema.

3 Conclusões

O estudo demonstrou que a topologia de uma única camada oculta com 8 neurônios foi a arquitetura mínima capaz de capturar a relação complexa entre as variáveis.

A análise da influência de diferentes funções de ativação mostrou que a escolha desse hiperparâmetro é crítica para o desempenho do modelo, mas que, quando bem-escolhida, a topologia é robusta. A função ReLU, apesar de popular, se mostrou inadequada para essa tarefa, falhando em modelar completamente a relação entre as variáveis. Por outro lado, as funções sigmoide e tangente hiperbólica (tanh) não apenas se ajustaram com perfeição aos dados, mas também garantiram que o treinamento fosse consistente, com múltiplos modelos convergindo para a mesma solução ideal.

A principal conclusão é que, para um desempenho consistente e confiável, a topologia de 8 neurônios combinada com uma função de ativação como a tanh ou a sigmoide é a escolha mais robusta. O experimento demonstrou que essas funções, ao contrário da ReLU, conseguem capturar a complexidade da curva e ofer-

ecem a confiabilidade necessária para que o modelo encontre a melhor solução em diferentes rodadas de treinamento