

**Relatório do Laboratório 7 - *Imitation Learning* com Keras**

## 1 Breve Explicação em Alto Nível da Implementação

O objetivo deste experimento laboratorial é projetar e implementar uma rede neural feedforward composta por três camadas densamente conectadas (fully connected layers), conforme especificado na Tabela 1, com o intuito de realizar aprendizado por imitação (imitation learning). A arquitetura da rede foi definida de modo a balancear capacidade de representação e eficiência computacional, utilizando funções de ativação não lineares apropriadas para promover a aprendizagem de padrões complexos.

Tabela 1: arquitetura da rede neural usada para o imitation learning.

Layer	Neurons	Activation Function
Dense	75	Leaky ReLU ( $\alpha = 0.01$ )
Dense	50	Leaky ReLU ( $\alpha = 0.01$ )
Dense	20	Linear

A implementação foi realizada por meio da biblioteca tensorflow.keras, uma API de alto nível para construção e treinamento de modelos de aprendizado profundo. O otimizador escolhido foi o Adam (Adaptive Moment Estimation), amplamente utilizado por sua robustez e capacidade de adaptação dinâmica da taxa de aprendizado para cada parâmetro da rede. Esse otimizador facilita o treinamento de redes profundas ao lidar eficientemente com gradientes ruidosos e sparse.

## 2 Figuras Comprovando Funcionamento do Código

### 2.1 Função de Classificação *sum\_gt\_zeros*

#### 2.2 $\lambda = 0$

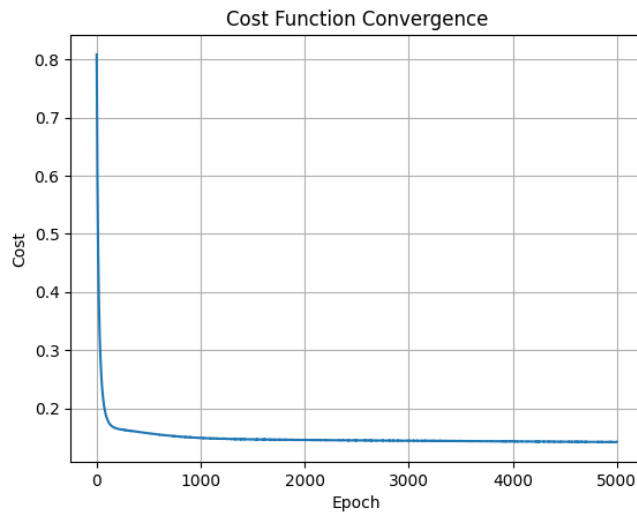


Figura 1: gráfico de convergência da função custo para a função *sum\_gt\_zeros*.

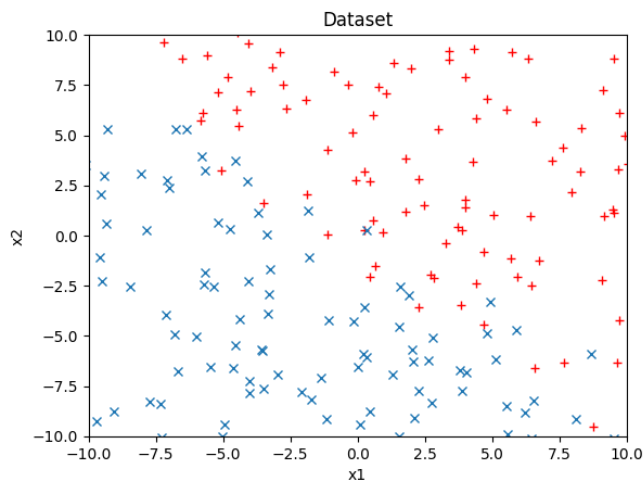


Figura 2: gráfico do dataset para a função *sum\_gt\_zeros*.

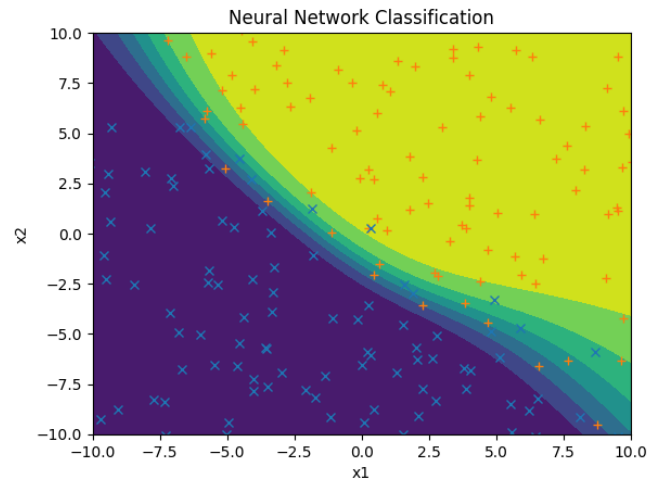


Figura 3: gráfico de classificação da rede neural para a função  $sum\_gt\_zeros$ .

### 2.3 $\lambda = 0.002$

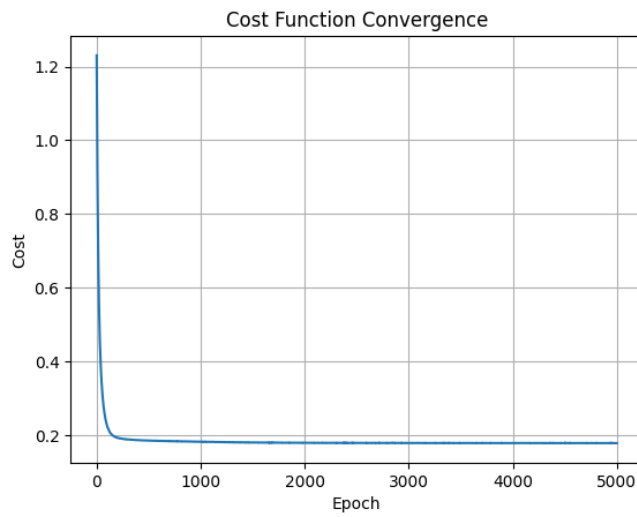


Figura 4: gráfico de convergência da função custo para a função  $sum\_gt\_zeros$ .

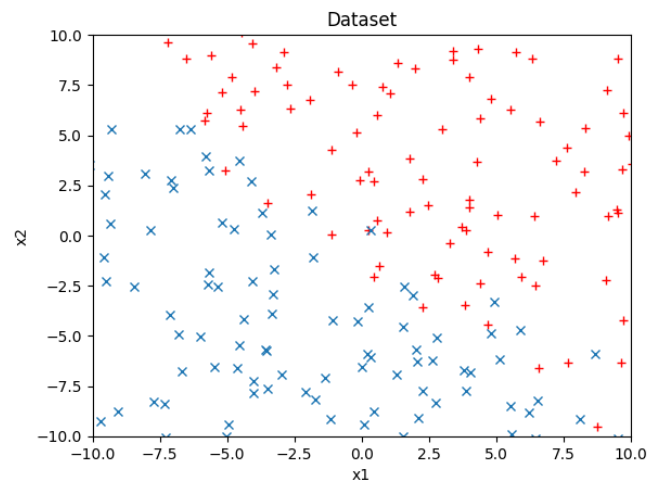


Figura 5: gráfico do dataset para a função  $sum\_gt\_zeros$ .

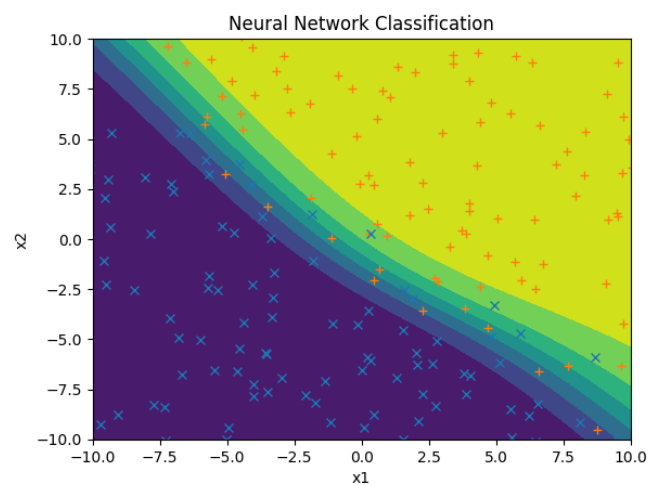


Figura 6: gráfico de classificação da rede neural para a função  $sum\_gt\_zeros$ .

## 2.4 Função de Classificação *XOR*

2.5  $\lambda = 0$

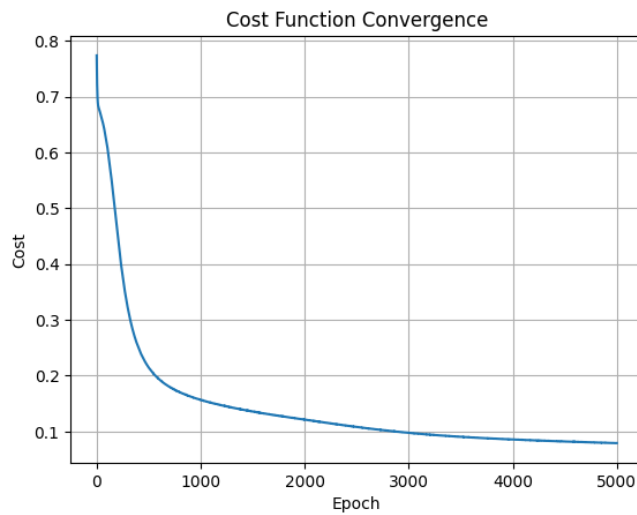


Figura 7: gráfico de convergência da função custo para a função *XOR*.

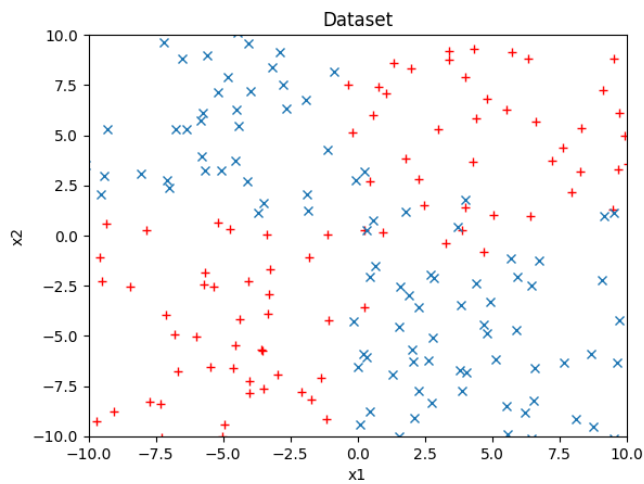


Figura 8: gráfico do dataset para a função *XOR*.

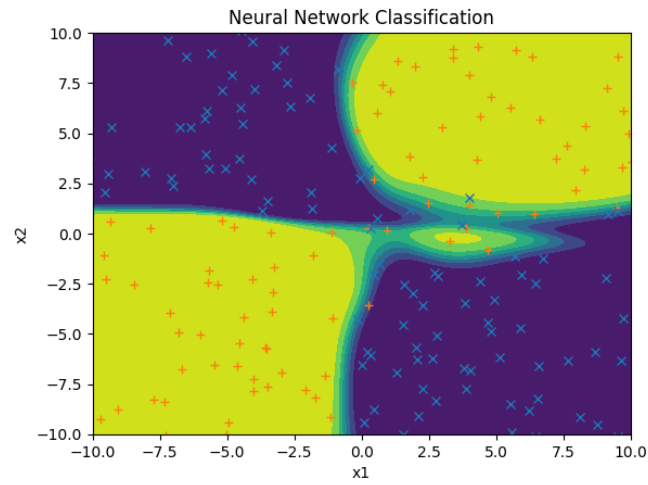


Figura 9: gráfico de classificação da rede neural para a função *XOR*.

**2.6**  $\lambda = 0.002$

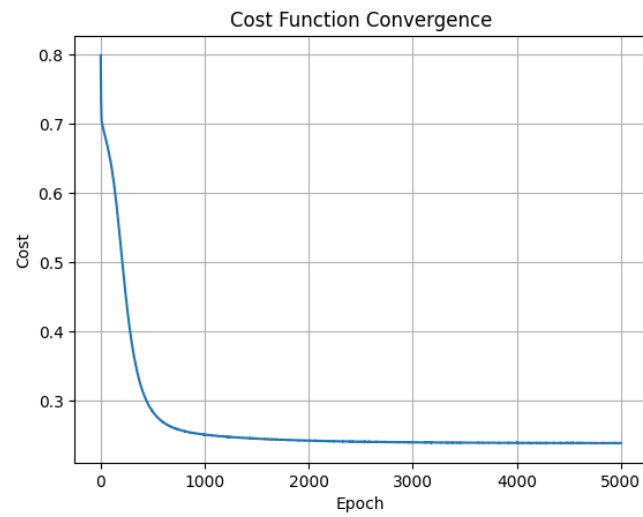


Figura 10: gráfico de convergência da função custo para a função *XOR*.

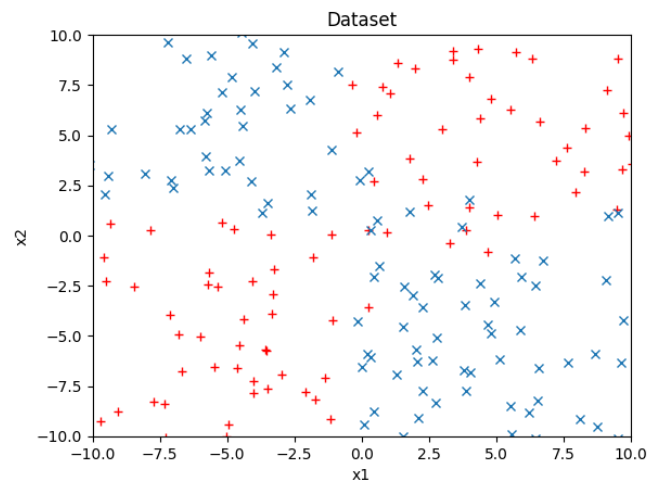


Figura 11: gráfico do dataset para a função  $XOR$ .

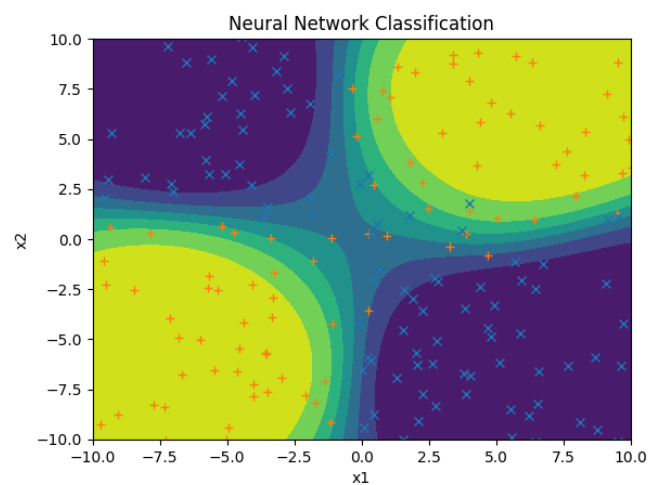


Figura 12: gráfico de classificação da rede neural para a função  $XOR$ .

## 2.7 Imitation Learning

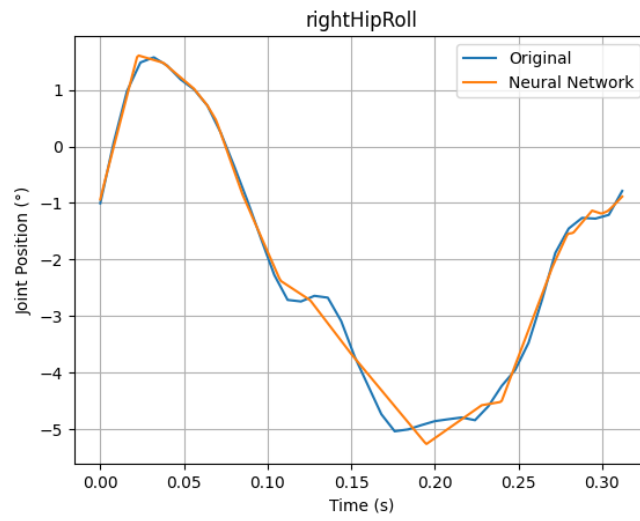


Figura 13: rightHipRoll gráfico no *Imitation Learning*.

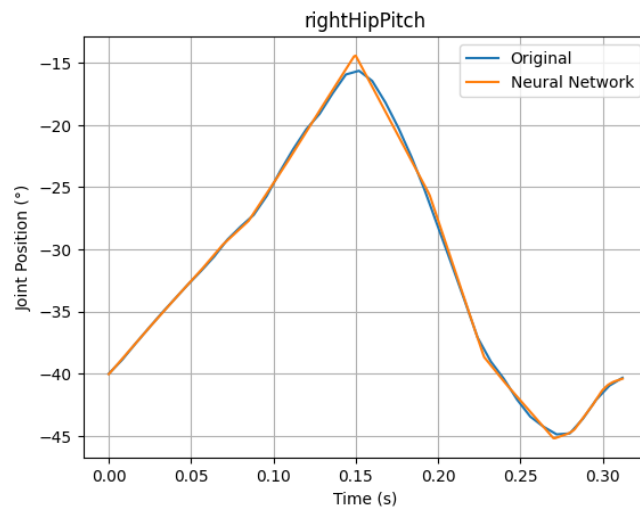


Figura 14: rightHipPitch gráfico no *Imitation Learning*.



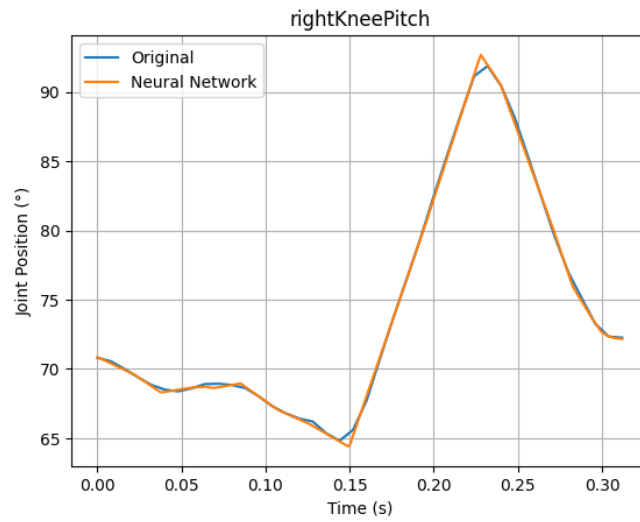


Figura 15: rightKneePitch gráfico no *Imitation Learning*.

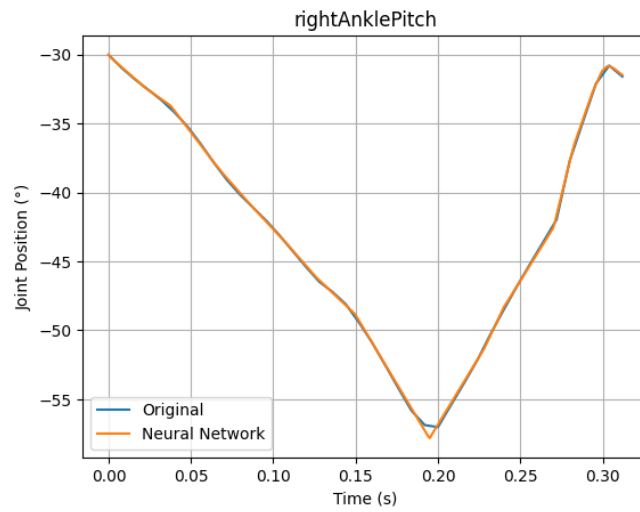


Figura 16: rightAnklePitch gráfico no *Imitation Learning*.

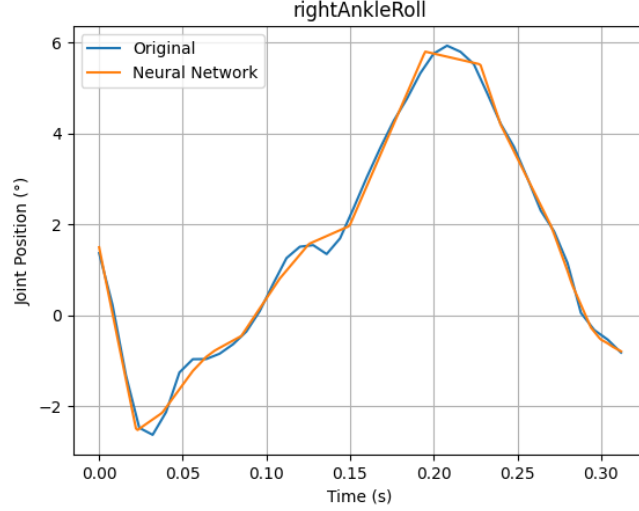


Figura 17: rightAnkleRoll gráfico no *Imitation Learning*.

### 3 Discussões

Durante a execução do laboratório, observou-se que o tempo de compilação e treinamento da rede neural foi elevado, mesmo para os casos de teste, como os problemas *sum\_gt\_zeros* e *XOR*. Nos experimentos de treinamento, foram analisados cenários com e sem regularização. Para  $\lambda = 0$ , a rede apresentou uma fronteira de decisão altamente flexível, com curvaturas acentuadas que sugerem sobreajuste aos dados de treinamento, incluindo ruídos e outliers, caracterizando um possível overfitting. Visualmente, a transição entre classes é irregular e excessivamente sensível aos exemplos.

Com a introdução da regularização  $L_2$   $\lambda = 0.002$ , a rede foi induzida a encontrar soluções mais simples, resultando em uma fronteira de decisão mais suave e estável. Isso evidencia um comportamento mais generalizável, com menor risco de sobreajuste. A regularização, portanto, mostrou-se eficaz na melhoria da robustez do modelo, contribuindo para um equilíbrio mais adequado entre viés e variância.

Para o *Imitation Learning*, observamos um comportamento muito próximo do esperado, com a rede conseguindo replicar de forma satisfatória a política do agente demonstrador ao longo dos episódios.