

INSTITUTO FEDERAL DO CEARÁ – IFCE
MESTRADO EM ENGENHARIA DE TELECOMUNICAÇÕES
REDES NEURAIS ARTIFICIAIS

GUILHERME LINCOLN DE LIMA TAVARES

TRABALHO 04
IMPLEMENTAÇÃO E ANÁLISE DE UMA REDE NEURAL
COM NEURÔNIOS SIGMOIDAIS

Professor: Ajalmar Rocha Neto, Dr.

Fortaleza – CE

2026

1 Introdução

Neste trabalho foi implementada e avaliada uma rede neural de uma camada de saída formada por neurônios sigmoidais. Foi utilizada a função logística como função de ativação, permitindo que cada neurônio produzisse uma saída contínua entre 0 e 1.

Os experimentos foram realizados com os conjuntos Iris e Artificial I. No conjunto Iris foram consideradas as três classes Setosa, Versicolor e Virginica, utilizando os quatro atributos disponíveis no treinamento principal. No conjunto Artificial I foram utilizadas três classes com 50 padrões cada.

Para representar as saídas desejadas foi utilizada a codificação 1-de- c (*1-out-of- c*). Para cada conjunto foram realizadas 20 execuções, sendo calculadas a acurácia média e o desvio padrão das taxas de acerto.

Também foram analisadas as superfícies de decisão do conjunto Artificial I e do conjunto Iris utilizando apenas dois atributos para permitir a visualização gráfica das regiões de classificação.

2 Bases de Dados Utilizadas

Foram utilizados dois conjuntos de dados: Iris e Artificial I.

2.1 Iris

O conjunto Iris foi utilizado considerando as três classes presentes no problema: Setosa, Versicolor e Virginica. No treinamento principal foram utilizados os quatro atributos do conjunto.

A rede possui três neurônios na camada de saída e as classes foram representadas pela codificação 1-de-3:

$$\text{Setosa} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \text{Versicolor} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{Virginica} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Quantidade de amostras: 150

Quantidade de atributos: 4

Quantidade de classes: 3

Para a visualização da superfície de decisão do Iris, foram utilizados apenas dois atributos: o comprimento da pétala e a largura da pétala.

2.2 Artificial I

O conjunto Artificial I foi formado por três classes, cada uma contendo 50 padrões, totalizando 150 amostras.

As saídas desejadas foram representadas por:

$$d_{\text{círculo}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad d_{\text{triângulo}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad d_{\text{estrela}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Quantidade de amostras: 150

Quantidade de atributos: 2

Quantidade de classes: 3

Parâmetros utilizados na geração: os três grupos foram gerados com 50 padrões por classe e desvio padrão igual a 0,85.

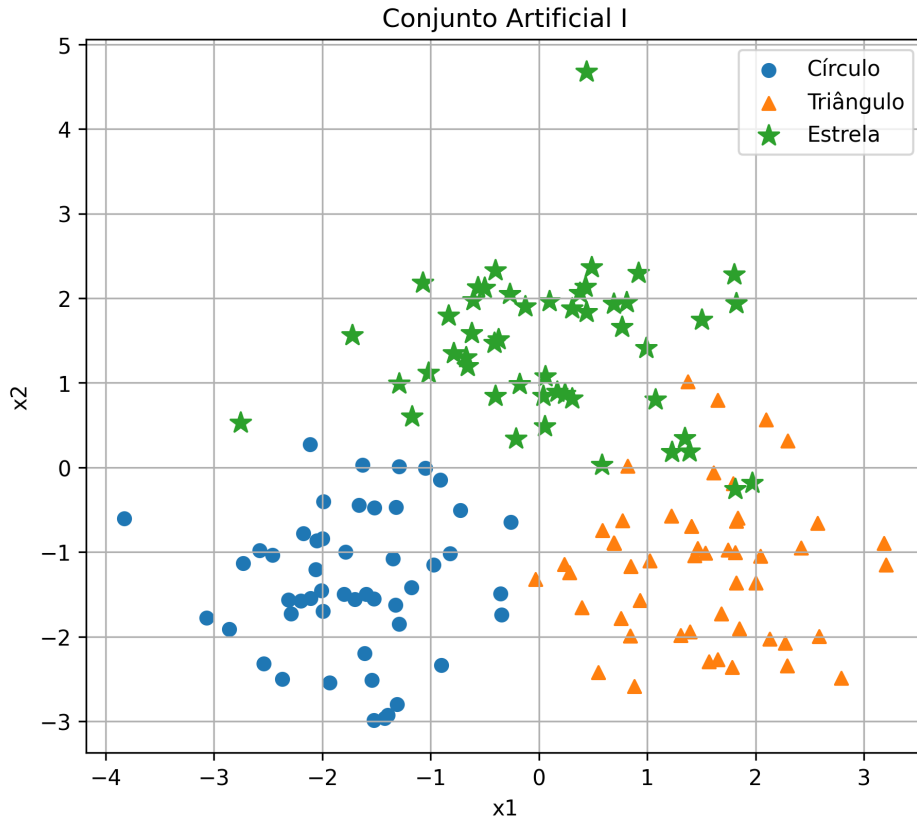


Figura 1: Conjunto de dados Artificial I utilizado nos experimentos.

3 Metodologia

3.1 Divisão dos Dados

Em cada realização, os dados foram divididos aleatoriamente em 80% para treinamento e 20% para teste. O conjunto de treinamento foi utilizado para ajustar os pesos da rede, enquanto o conjunto de teste foi utilizado para avaliar seu desempenho.

Os atributos foram padronizados em cada realização. Os parâmetros da padronização foram obtidos utilizando apenas o conjunto de treinamento e posteriormente aplicados ao conjunto de teste.

3.2 Número de Realizações

Para cada conjunto de dados foram realizadas 20 execuções independentes. Em cada realização foi utilizada uma nova divisão dos dados em treinamento e teste.

Ao final das 20 realizações, foram calculadas a acurácia média e o desvio padrão das taxas de acerto.

3.3 Rede Neural com Neurônios Sigmoidais

A rede utilizada possui uma única camada de saída composta por c neurônios, em que c representa a quantidade de classes do problema.

Para cada neurônio de saída j , a soma ponderada foi calculada por:

$$u_j = \sum_{i=1}^n w_{ji}x_i + b_j \quad (1)$$

em que w_{ji} representa o peso associado à entrada x_i no neurônio j , e b_j representa o bias desse neurônio.

Foi utilizada a função logística como função de ativação:

$$y_j = \frac{1}{1 + e^{-u_j}} \quad (2)$$

O erro de cada neurônio foi calculado por:

$$e_j = d_j - y_j \quad (3)$$

em que d_j representa a saída desejada e y_j representa a saída calculada pelo neurônio.

Considerando a derivada da função logística, os pesos foram atualizados pela regra:

$$w_{ji}(t+1) = w_{ji}(t) + \alpha e_j y_j (1 - y_j) x_i \quad (4)$$

e o bias foi atualizado por:

$$b_j(t+1) = b_j(t) + \alpha e_j y_j (1 - y_j) \quad (5)$$

em que α representa a taxa de aprendizagem.

Taxa de aprendizagem utilizada:

$$\alpha = 0.01$$

Número máximo de épocas: 500

3.4 Codificação 1-de- c

Para os problemas com três classes, a saída desejada foi representada por um vetor com três posições. Apenas a posição correspondente à classe correta recebeu valor 1, enquanto as demais receberam valor 0.

$$d_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad d_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad d_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Durante a predição, a classe associada ao maior valor de saída entre os neurônios foi considerada como a resposta da rede.

3.5 Métrica de Avaliação

A acurácia foi calculada pela proporção de amostras de teste classificadas corretamente:

$$\text{Acurácia} = \frac{\text{Número de classificações corretas}}{\text{Número total de amostras de teste}} \quad (6)$$

Após as 20 realizações, foram calculadas a acurácia média e o desvio padrão das taxas de acerto obtidas em cada conjunto.

4 Resultados Experimentais

Os resultados das 20 realizações são apresentados por meio da acurácia média e do desvio padrão para os conjuntos Iris e Artificial I.

Tabela 1: Resultados obtidos nas 20 realizações.

Conjunto	Acurácia média (%)	Desvio padrão (%)
Iris	92,83	5,70
Artificial I	93,33	3,50

Os dois conjuntos apresentaram acurácias médias semelhantes. O conjunto Iris alcançou **92,83%**, com desvio padrão de **5,70%**, enquanto o Artificial I apresentou acurácia média de **93,33%**, com desvio padrão de **3,50%**.

O Artificial I apresentou menor variação entre as 20 realizações, enquanto o Iris apresentou maior diferença de desempenho entre as diferentes divisões dos dados em treinamento e teste.

5 Superfície de Decisão – Artificial I

Para o conjunto Artificial I, foi selecionada a **realização 1** como representativa, pois apresentou acurácia de **93,33%**, valor igual à acurácia média obtida nas 20 execuções.

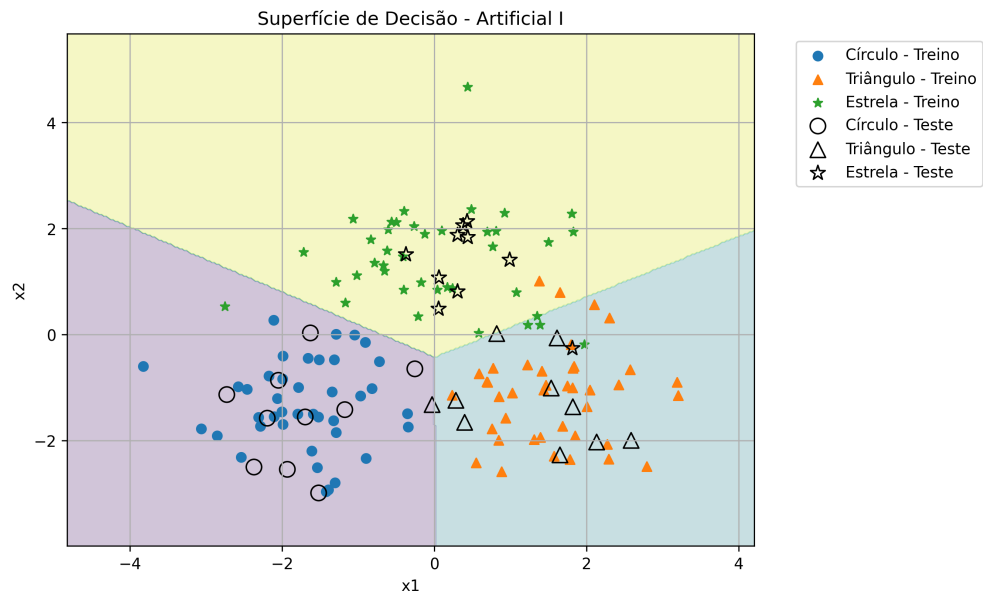


Figura 2: Superfície de decisão obtida pela rede sigmoidal para o conjunto Artificial I.

Realização selecionada: 1

Acurácia da realização: 93,33%

Análise do gráfico: a superfície de decisão apresenta três regiões distintas, correspondentes às classes Círculo, Triângulo e Estrela. De modo geral, observa-se que a rede conseguiu separar adequadamente os três grupos, embora existam pontos localizados próximos às fronteiras entre as regiões. Esse comportamento é coerente com a acurácia de 93,33%, indicando que o modelo obteve bom desempenho, mas sem separação perfeita entre todas as amostras.

6 Superfície de Decisão – Iris com Dois Atributos

Para permitir a visualização da superfície de decisão do conjunto Iris, foi realizado um experimento adicional utilizando apenas dois atributos: **comprimento da pétala** e **largura da pétala**. Nesse experimento também foram realizadas 20 execuções. Essa representação foi utilizada apenas para a análise gráfica, enquanto o experimento principal com o Iris utilizou os quatro atributos.

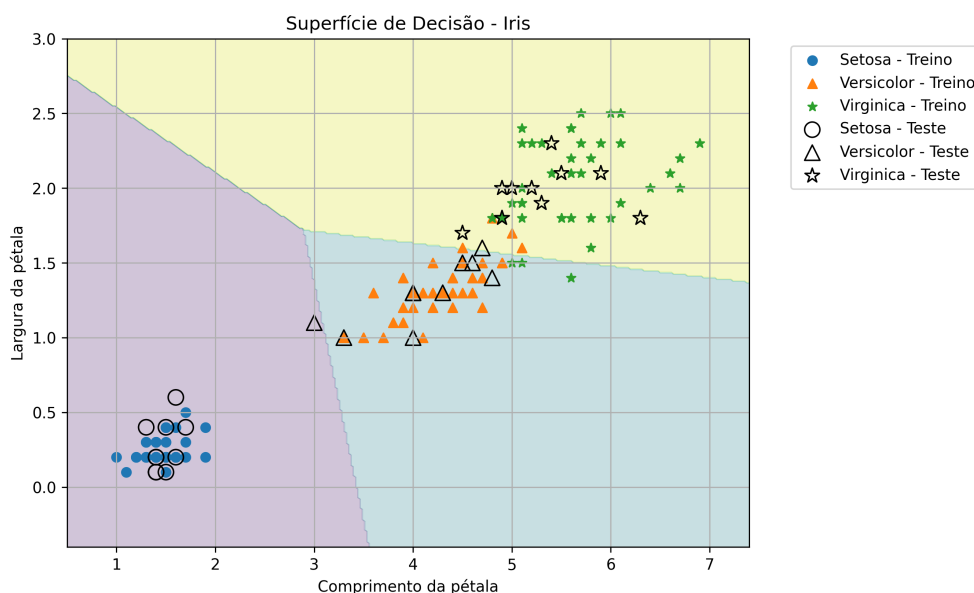


Figura 3: Superfície de decisão obtida para o conjunto Iris utilizando dois atributos.

Atributos utilizados: comprimento da pétala e largura da pétala

Realização selecionada: 2

Acurácia da realização: 93,33%

Acurácia média: 94,50%

Análise do gráfico: a superfície de decisão mostra três regiões de classificação associadas às classes Setosa, Versicolor e Virginica. Observa-se que a classe Setosa ficou bem separada das demais, enquanto Versicolor e Virginica apresentaram regiões mais próximas, com alguns pontos localizados próximos às fronteiras de decisão. Esse comportamento é coerente com a acurácia média de 94,50%, indicando que a rede obteve bom desempenho mesmo utilizando apenas dois atributos.

7 Discussão dos Resultados

Os resultados mostram que a rede sigmoidal conseguiu um bom desempenho nos dois conjuntos utilizados.

Iris: utilizando os quatro atributos, a acurácia média foi de **92,83%**, com desvio padrão de **5,70%**. Entre os experimentos principais, foi o conjunto que apresentou a maior variação entre as 20 realizações.

Artificial I: a rede obteve acurácia média de **93,33%** e desvio padrão de **3,50%**. O resultado ficou próximo do Iris, mas com uma variação menor entre as realizações. Pelo gráfico da superfície de decisão, também foi possível observar que a maior parte das amostras ficou dentro da região correspondente à sua classe, apesar de alguns pontos estarem mais próximos das regiões das outras classes.

Iris com dois atributos: utilizando apenas o comprimento e a largura da pétala, a acurácia média foi de **94,50%**, com desvio padrão de **3,69%**. Mesmo utilizando somente dois atributos, o resultado ficou próximo do experimento com os quatro atributos. No

gráfico, a classe Setosa ficou bem separada, enquanto Versicolor e Virginica ficaram mais próximas entre si.

De forma geral, a rede apresentou resultados próximos nos experimentos. Os gráficos também ajudaram a visualizar onde ocorreram as maiores dificuldades de separação entre as classes.

8 Conclusão

Neste trabalho foi implementada uma rede neural com uma camada de saída formada por neurônios sigmoidais, utilizando a função logística e a codificação 1-de-*c* para representar as classes.

A rede foi testada nos conjuntos Iris e Artificial I, realizando 20 execuções para cada caso. No Iris, utilizando os quatro atributos, foi obtida acurácia média de **92,83%**. No Artificial I, a acurácia média foi de **93,33%**.

Também foram geradas as superfícies de decisão do Artificial I e do Iris utilizando dois atributos. Esses gráficos ajudaram a visualizar como a rede separou as diferentes classes e onde ficaram os pontos mais próximos das regiões de decisão.

De forma geral, os resultados mostraram que a rede sigmoidal conseguiu realizar a classificação dos dois conjuntos com bom desempenho, mantendo resultados próximos ao longo das diferentes realizações.