

**INSTITUTO FEDERAL DE EDUCAÇÃO
CIÊNCIA E TECNOLOGIA DE MINAS GERAIS -
CAMPUS BAMBUÍ**

**BACHARELADO EM ENGENHARIA DE
COMPUTAÇÃO**

GABRIELA DÂMASO REZENDE

TÓPICOS ESPECIAIS EM ENGENHARIA DE COMPUTAÇÃO

RELATÓRIO 4

BAMBUÍ - MG

2023

1 Questão 2

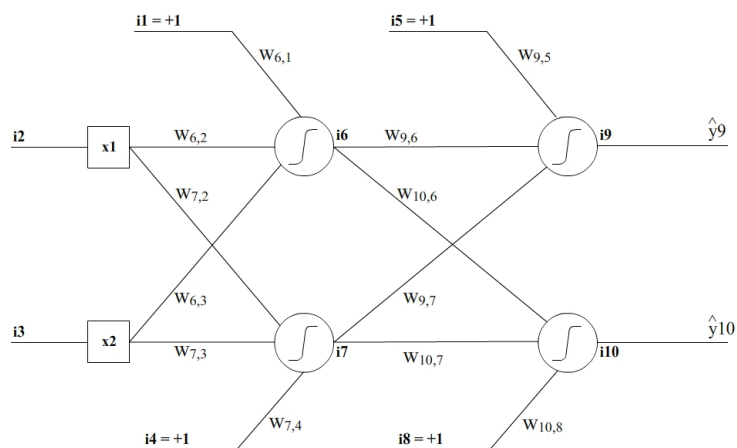


Figura 1: Esquema da rede

Na figura 1 temos o esquema da rede com 2 entradas e 2 saídas, as entradas são os valores de x_1 e x_2 que são aplicados nos neurônios da segunda camada e multiplicados por um certo peso e bias. Cada saída na segunda camada sofre o mesmo processo, é multiplicada por um peso e um bias e calculado em um neurônio. E os valores de saída dessa camada são os valores de i_9 e i_{10} que correspondem a saída da rede.

No código temos essa mesma ideia da figura, onde temos as entradas os pesos e bias, e a saída da rede é calculada com base nesses valores.

2 Questão 3

```
[1] "Epoca 9990"
      x1 x2      yhat1      yhat2
[1,] -1  1 -0.9621168  0.9628803
[2,]  1 -1  0.9529898 -0.9539039
[3,]  1 -1  0.9525225 -0.9534646
[4,] -1  1 -0.9669982  0.9676889
> |
```

Figura 2: Matriz de saídas

A matriz representada na figura 2 mostra os valores de x_1 e x_2 e o quão próximos os valores \hat{y}_1 e \hat{y}_2 chegaram deles.

3 Questão 4

3.1 Taxa de aprendizado

A taxa de aprendizado é um valor que define o quão rápido a rede neural vai aprender, se o valor for muito alto a rede pode não aprender (pode ficar "sambando" entre picos e vales), e se for muito baixo a rede pode demorar muito para aprender (ao andar lentamente de um ponto a outro). A seguir estão alguns exemplos de gráficos que tiveram uma variação na taxa de aprendizado.

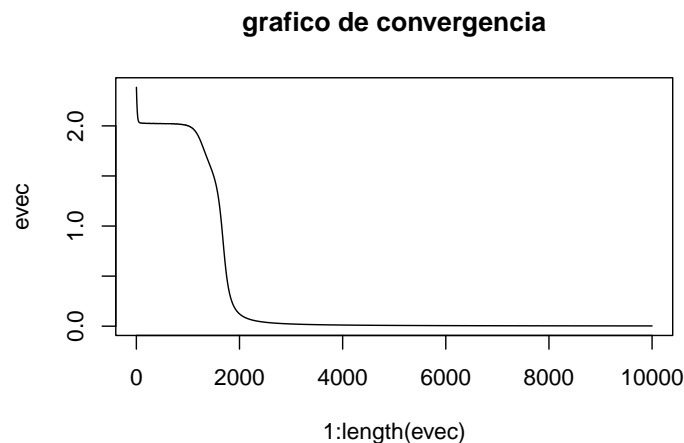


Figura 3: Taxa de aprendizado = 0.01

Na figura 3 temos a taxa de aprendizado igual a 0.01, que é a taxa original dada no algoritmo, podemos ver que a rede aprendeu bem ao chegar a um valor de erro mínimo.

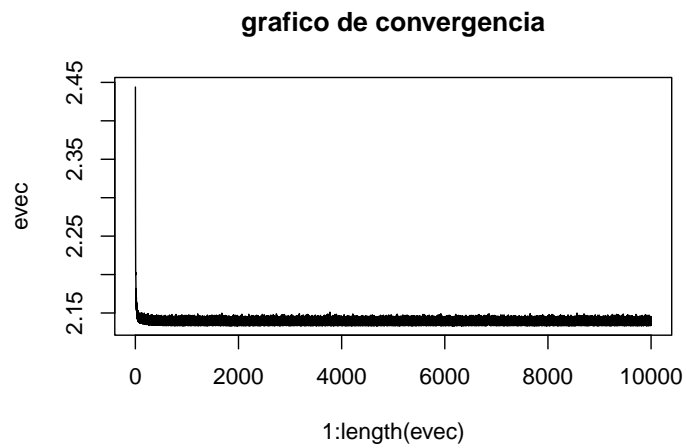


Figura 4: Taxa de aprendizado = 0.07

Na figura 4 temos a taxa de aprendizado igual a 0.07, podemos ver que a rede apresentou ruído em seu gráfico.

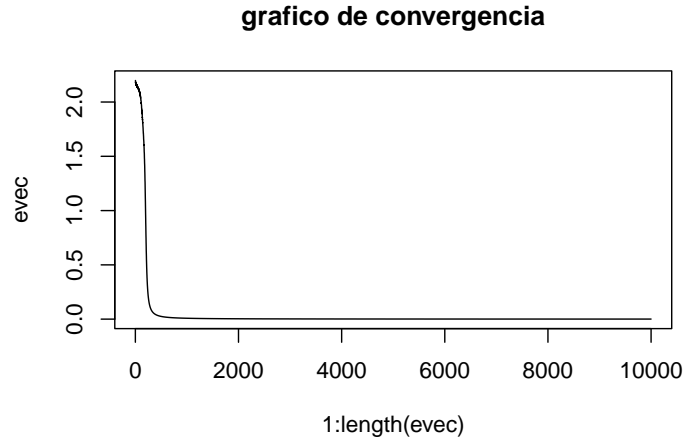


Figura 5: Taxa de aprendizado = 0.04

Na figura 5 temos a taxa de aprendizado igual a 0.04, podemos ver que a rede aprendeu pois chegou a um número próximo de 0, e demorou ainda menos épocas que a rede com taxa de aprendizado igual a 0.01.

3.2 Máximo de épocas

O número de épocas é o número de vezes que a rede neural vai passar por todos os dados de treinamento, ou seja, o número máximo de épocas é o limite de vezes que teremos o laço de repetições que a rede neural faz.

Aumentar o máximo de épocas permite que a MLP tenha mais oportunidades de ajustar seus pesos e melhorar seu desempenho. No entanto, se o máximo de épocas for muito alto, a rede pode acabar sobreajustando os dados, o que resulta em um desempenho inferior. Por outro lado, se o máximo de épocas for muito baixo, a rede pode não ter tempo suficiente para convergir para uma solução adequada.

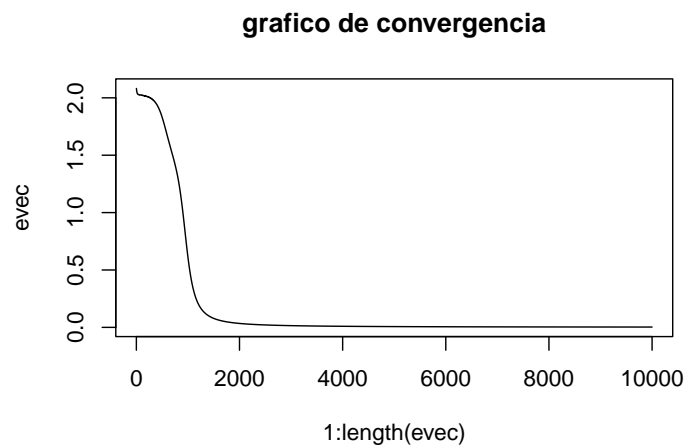


Figura 6: Máximo de épocas = 10000

Na figura 6 temos o máximo de épocas igual a 10000, que foi o valor original apresentado no código. Ademais, podemos ver que a rede aprendeu bem ao chegar a um valor de erro mínimo.

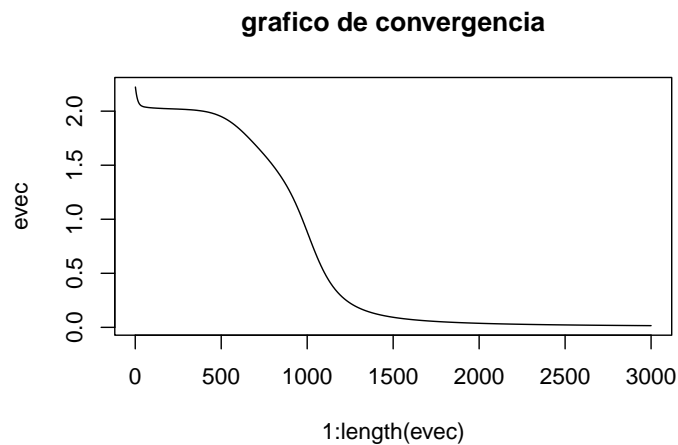


Figura 7: Máximo de épocas = 3000

Na figura 7 temos o máximo de épocas igual a 3000, podemos ver que a rede conseguiu convergir apesar do número de épocas ter diminuído em 7000 em relação a figura 6.

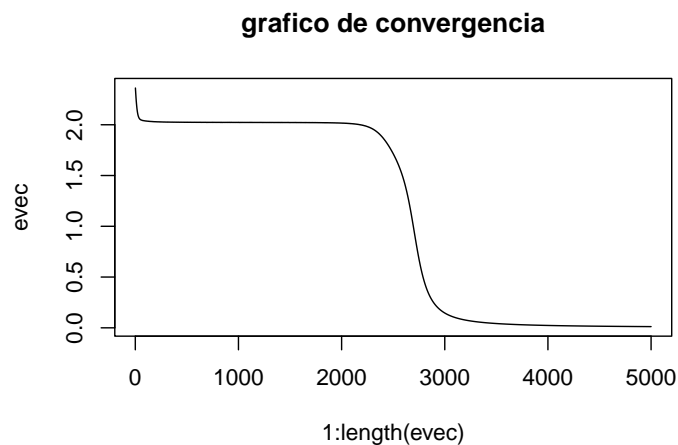


Figura 8: Máximo de épocas = 5000

Na figura 8 temos o máximo de épocas igual a 5000, podemos ver que a rede conseguiu convergir apesar do número de épocas ter diminuído em 5000 em relação a figura 6. Porém, em relação a figura 7 a rede demorou mais para convergir.

3.3 Tolerância ao erro

A tolerância ao erro é uma medida que determina quando o treinamento deve parar com base na diferença entre a saída da rede neural e o valor esperado para cada exemplo de treinamento. Se a diferença for menor que a tolerância ao erro, o treinamento é considerado concluído.

Aumentar a tolerância ao erro permite que o treinamento pare mais cedo, o que pode economizar tempo de computação. No entanto, uma tolerância ao erro muito alta pode resultar em uma rede que não está suficientemente ajustada aos dados de treinamento, levando a um desempenho inferior.

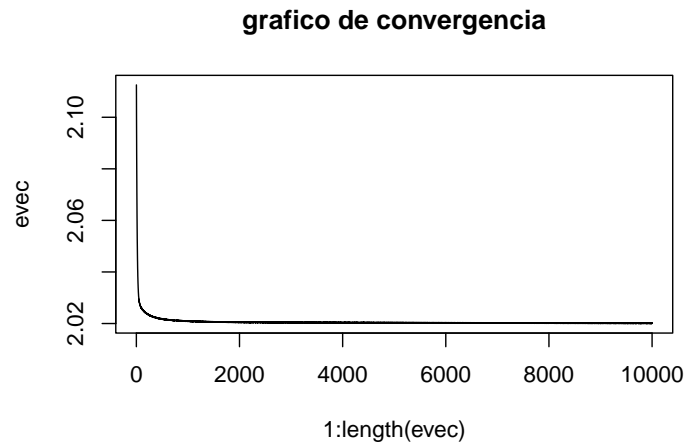


Figura 9: Tolerância ao erro = 0

Na figura 9 temos a tolerância ao erro igual a 0, que foi o valor original apresentado no código.

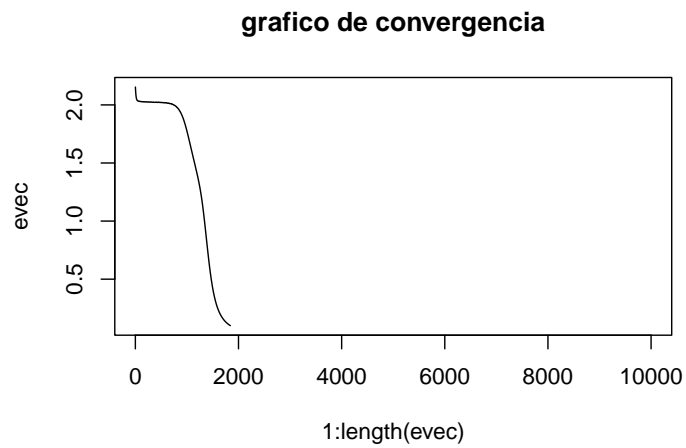


Figura 10: Tolerância ao erro = 0.1

Na figura 10 temos a tolerância ao erro igual a 0.1. O gráfico apresentado por tolerâncias diferentes de 0 são muito parecidos e de certa maneira incompletos como também podemos observar na figura 11 onde a tolerância é de 1.

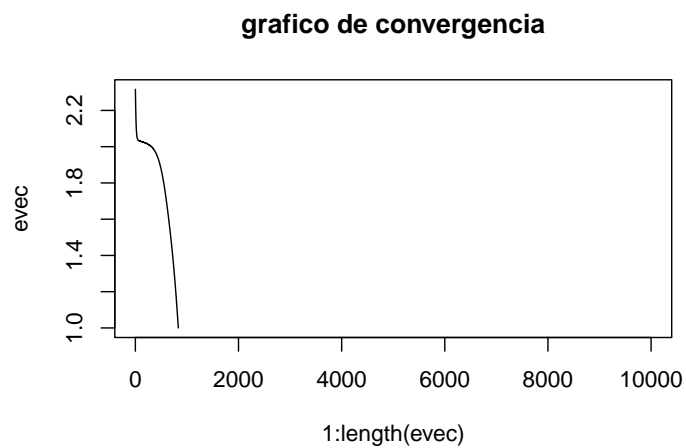


Figura 11: Tolerância ao erro = 1

4 Questão 5

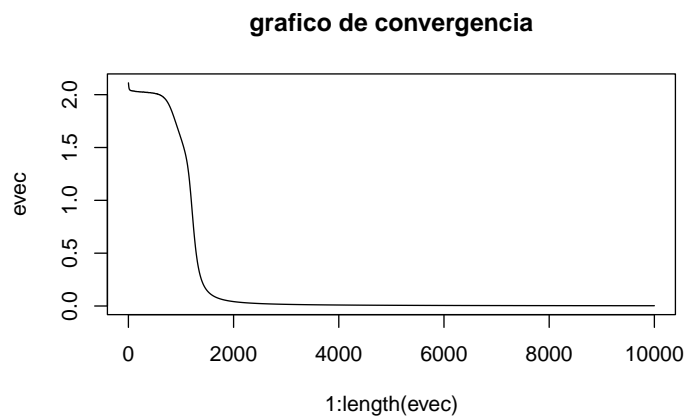


Figura 12: Gráfico da questão 5

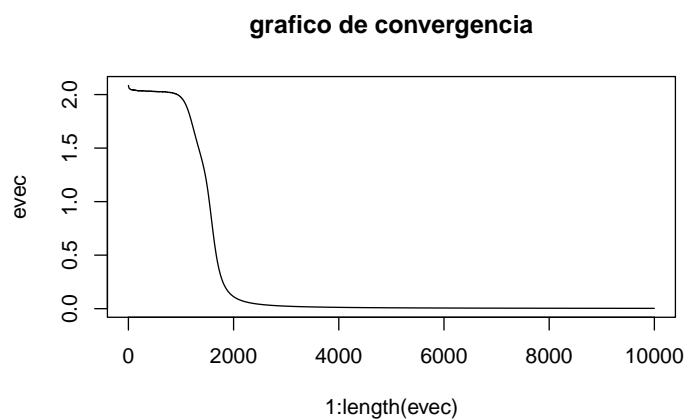


Figura 13: Gráfico da questão 5

Apesar de muito parecidos os gráficos 11 e 12 não são iguais. Mesmo tendo os mesmos valores de entrada, de tolerância e de épocas.

Isso acontece porque a rede neural é inicializada com pesos aleatórios, e isso faz com que a rede neural tenha um comportamento diferente a cada vez que é executada.

Porém, pode acontecer de os gráfico serem iguais mesmo em execuções diferentes. Um exemplo são as figuras 14 e 15 que podemos afirmar que são

iguais as figuras 12 e 13 respectivamente mesmo com pesos aleatórios e em execuções diferentes.

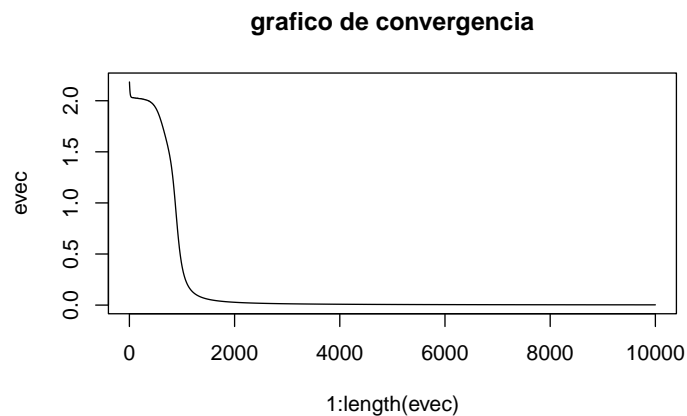


Figura 14: Gráfico da questão 5

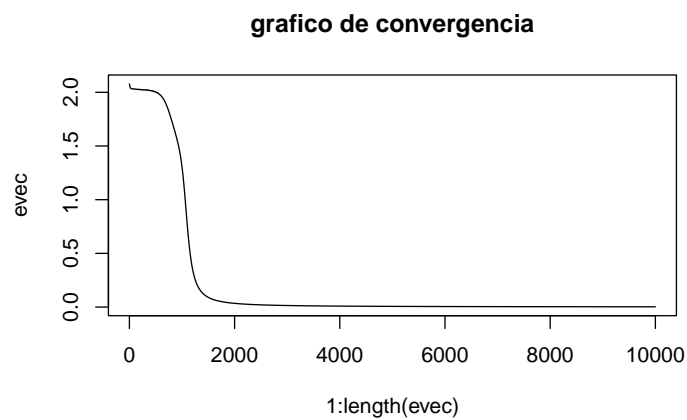


Figura 15: Gráfico da questão 5

5 Questão 6

A Perceptron e a Adaline são modelos de redes neurais artificiais mais simples em comparação com a MLP.

O Perceptron é um modelo de rede neural de uma única camada, composto por um único neurônio de saída. Ele é projetado para realizar classifi-

cação binária, onde o objetivo é separar dois conjuntos de dados. Portanto acredito que o Perceptron não é capaz de resolver o problema proposto da mesma maneira.

A Adaline também é um modelo de rede neural de uma única camada. Ela utiliza uma função de ativação linear e usa o algoritmo de descida do gradiente para ajustar os pesos sinápticos. Ela é capaz de lidar com problemas linearmente separáveis. Portanto, acredito que ela também não é capaz de resolver o problema proposto da mesma maneira.