Redes Neurais

Augusto Rodrigo Camblor Santos

02/06/ 2023

1 Descrição

Este projeto tem como base um estudo sobre redes neurais, onde será utilizado um código base fornecido em aula pelo professor ministrador do curso de cálculo.

Neste projeto, a finalidade é um estudo prático onde ocorre uma classificação através de entradas no console; as três primeiras entradas correspondem às entradas dos pesos denominados como W, na sequência há mais três entradas que correspondem ao bias denominadas como B, por fim as duas últimas entradas são determinadas por *learning rate* (lr) e a taxa de aprendizagem de erro (err). A partir disso, há uma captação dos valores do gradiente que resultará em um treinamento para classificar se um número é primo ou não, e se ele é par ou ímpar.

2 O que eu quero Classificar

Número Pares e Ímpares

No livro de *Fundamentos da Matemática Elementar* volume 1 do Gelson Iezzi:

> "Quando a é divisor de b, dizemos que 'b é divisível por a' ou 'b é múltiplo de a'."
>
> "Para um inteiro a qualquer, indicamos com $D(a)$ o conjunto de seus divisores e com $M(a)$ o conjunto de seus múltiplos."(Iezzi, p. 43)[1]

Com esse fundamento, a classificação dos números pares ou ímpares é determinada pela divisibilidade por 2. Segundo Carl Friedrich Gauss: "Um número é par se for divisível por 2, e ímpar caso contrário". Portanto, ao realizar a divisão e o número do resultado for um número inteiro, a classificação é par; caso contrário, a classificação é ímpar. E de acordo com o matemático francês Évariste Galois: "Os números pares são como soldados, marchando em pares e facilmente alinhados. Os números ímpares são como artistas temperamentais, nunca se encaixando perfeitamente".

Exemplos dos números pares:

- $4 \div 2 = 2$ (resultado inteiro, portanto 4 é par)
- $10 \div 2 = 5$ (resultado inteiro, portanto 10 é par)

Exemplos dos números ímpares:

- $7 \div 2 = 3.5$ (resultado não inteiro, portanto 7 é ímpar)
- $15 \div 2 = 7.5$ (resultado não inteiro, portanto 15 é ímpar)

A expressão algébrica que comprova as afirmações e citações anteriores:

- $2n + 1$ para números ímpares
- $2n$ para números pares

Em suma, a divisibilidade por 2 é o critério principal para classificar um número como par ou ímpar. Essa distinção é essencial na estruturação dos números inteiros, permitindo a compreensão e manipulação dos padrões numéricos. Portanto, no código feito para o treinamento, não valida números negativos por não fazerem parte do conjunto, sendo assim, a resposta retornada no console para casos numéricos negativos será "False".

Números Primos

Um número é considerado primo se ele possui exatamente dois divisores, sendo o 1 e ele mesmo. Em outras palavras, um número primo é um número inteiro maior que 1 que não pode ser dividido de forma exata por nenhum outro número além de ele mesmo e pelo 1.

Outro meio de encontrar um número primo é utilizando o método da divisão por todos os números inteiros menores que ele mesmo, até a raiz quadrada desse número. Se nenhum desses números for um divisor exato, então o número é primo. Com exceção do número 2, por ser o único número par que é primo, é possível fazer uma verificação separada para ele e, em seguida, aplicar o método da divisão para os demais números ímpares.

Segundo Gauss, "Os números primos são os indivíduos solitários do reino dos números, sem divisores além de si mesmos e da unidade". Ressaltando que a verificação dessa classificação se torna mais eficiente para números grandes utilizando algoritmos mais avançados, como Crivo de Eratóstenes ou o teste de primalidade de Miller-Rabin.

No livro de *Fundamentos da Matemática Elementar* volume 1 do Gelson Iezzi:

> "Uma importante noção que devemos ter sobre números inteiros é o conceito de divisor".

> "Dizemos que o inteiro a é divisor do inteiro b (símbolo $a|b$) quando existe um inteiro c tal que $ca = b$." (Iezzi, p. 43)[1]

Nos trazendo a expressão algébrica:

$a \,|\, b \Leftrightarrow (\,\exists\; c\, \in \mathbb{Z} \,|\, ca = b).$

Exemplos de números primos e seus casos:

- O número 2 é o único número primo par, pois só é divisível por 1 e 2.
- O número 3 também é primo, pois só é divisível por 1 e 3.
- O número 4 não é primo, pois além de ser divisível por 1 e 4, também é divisível por 2.

Em resumo, os números primos são elementos fundamentais na estrutura dos números e possuem propriedades únicas que os diferenciam dos demais.

3 Função de Erro

Abaixo uma imagem da qual representa a função erro e a equação y feitas a mão, dentro do código a função erro pode ser encontrada como 'errnovo' dentro da função ' *def descentV* '.

A tolerância feita no código foi de: 10^{-6}, encontrada no código como '*tol*' dentro ' *descentV* ' também.

No código a equação $\hat{y}_i$ pode ser encontrada na função ' *predict* ' em seu retorno.

$$\hat{y}_i = \sigma(w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 * x + b_1) + b_2) + b_3)$$

$$\mathcal{E}(w_1, w_2, w_3, b_1, b_2, b_3) = 1/2 \sum_{i=1}^{n} (y - \hat{y}_i)^2$$

Figure 1: Ilustra a função "erro" e a equação $\hat{y}_i$ feitas a mão

```
Último erro quadrático:  6.8514732502072990-08
```

Figure 2: Último erro quadrático

4 Sigmoide

Abaixo seguem dois exemplos de sigmoide. O primeiro ilustra como a função geradora do gráfico está funcionando com valores que geram grau de significância para a sigmoide. Já o segundo exemplo é uma 'plotagem' da sigmoide com os valores utilizados como exemplos citados mais abaixo no relatório. A função referida no código é encontrada em duas etapas.

A função que realiza o cáculo se chama ' *sigmoid* '. Já a função que realiza a 'plotagem' do gráfico se chama '*plot_sigmoid*'.

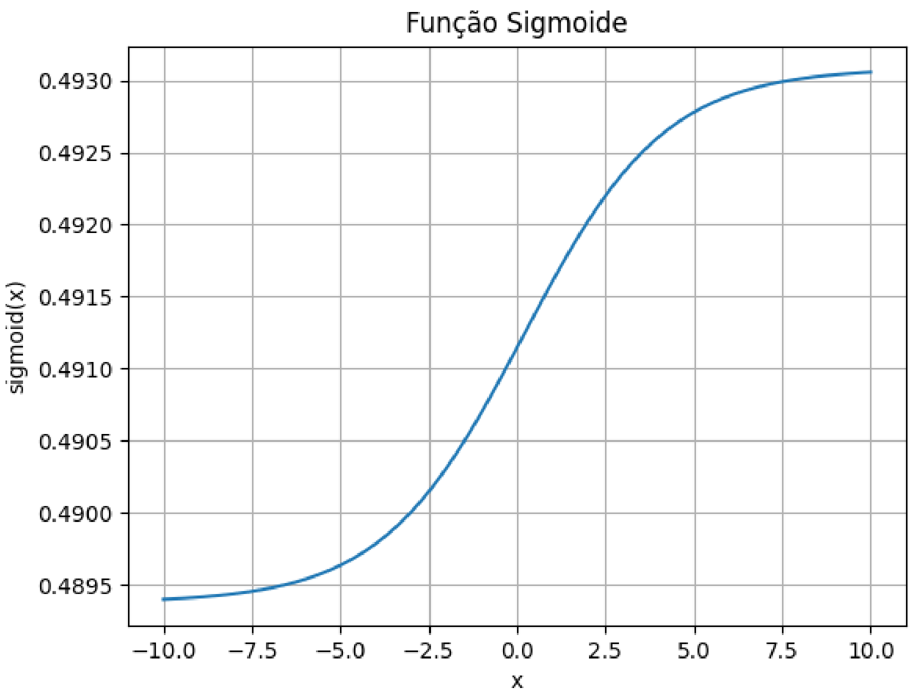

Figure 3: Sigmoide com valores de pesos e bias definidos

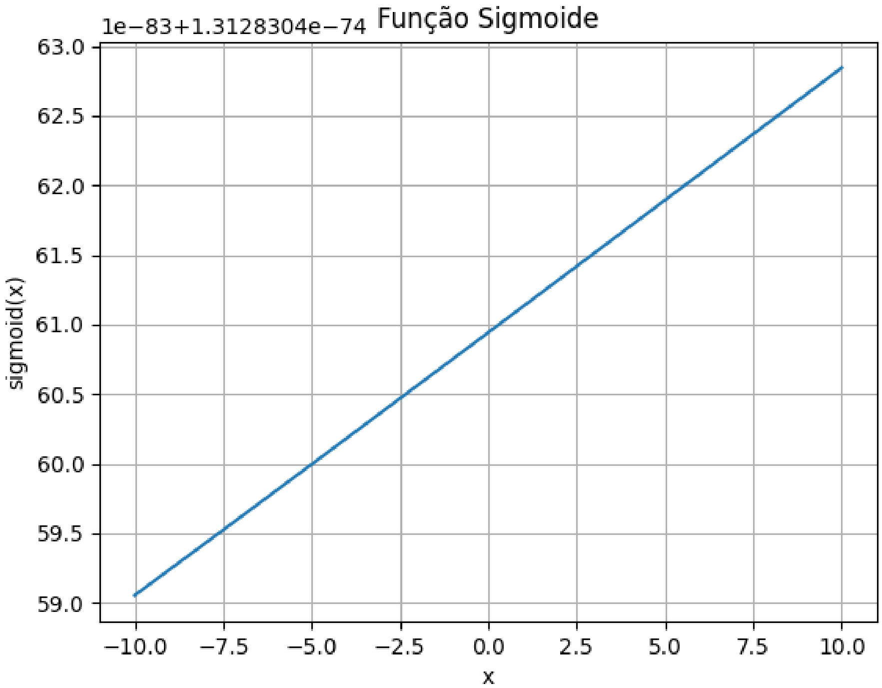

Figure 4: Sigmoide com valores não definidos

5 Gradiente

Abaixo há a plotagem do gráfico do gradiente dos valores exemplos. A função no código pode ser encontrada em duas etapas. A que realiza o cálculo se chama ' *descentV* '. Já a função que realiza a criação do gráfico se chama '*plot_gradient*'.

A função que realiza o gráfico é acompanhada das e chamada após o cálculo feita na função ' *neural* que utiliza as derivadas da função $\hat{y}_{\mathrm{i}}$.

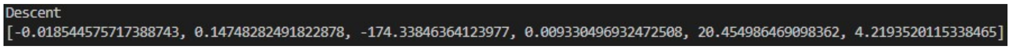
```
Descent
[-0.018544575717388743, 0.14748282491822878, -174.33846364123977, 0.009330496932472508, 20.454986469098362, 4.2193520115338465]
```

Figure 5: Valores do gradiente

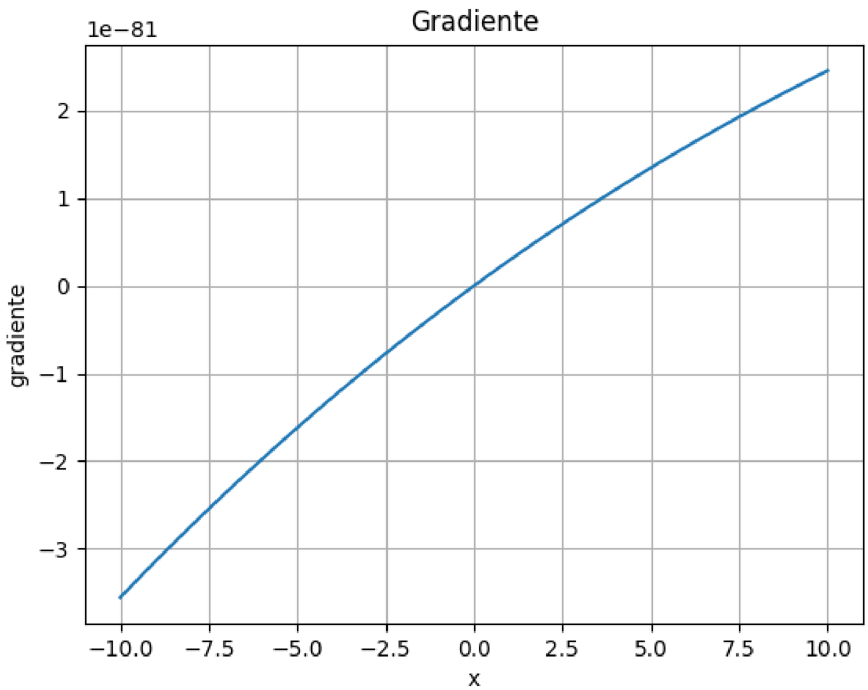

Figure 6: Gradiente dos valores informados

5.1 Derivadas dos pesos e bias feitas a mão

Calculo das derividas dos pesos feitos a mão

$$\frac{\partial \ell}{\partial w_1} = 0$$

$$\frac{\partial \ell}{\partial w_1} = \left[1/2 \, (y - \underbrace{\sigma(w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 * x_i + b_1) + b_2) + b_3))}_{\hat{y}_i} \right]^2 = (y - \hat{y}_i)(y - \hat{y}_i)_{w_1}$$

2

$$= -(y - \hat{y}_i)\left(\sigma.(w_3 . \sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2) + b_3)\right)_{w_1}$$

$$= -(y - \hat{y}_i)\,\sigma'(w_3 . \sigma(w_2 . \sigma(w_1 \cdot x_i + b_1) + b_2) + b_3) . \left[w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 \cdot x + b_1) + b_2) + b_3 \right]_{w_1}$$

$$= -(y - \hat{y}_i)\,\hat{y}_i (1 - \hat{y}) . w_3 \left[\sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2) \right]_{w_1}$$

$$= -(y - \hat{y}_i)\,\hat{y}_i (1 - \hat{y}) . w_3 \cdot w_2 \cdot \sigma(w_1 \cdot x_i + b_1) \cdot (1 - \sigma(w_1 \cdot x_i + b_1)) \cdot x_i$$

Figure 7: Derivada w1

$$\frac{\partial \varepsilon}{\partial w_2} = \left[1/2 \, (y - \underbrace{\sigma(w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 * x_i + b_1) + b_2) + b_3))}_{\hat{y}_i} \right]^2 = (y - \hat{y}_i)(y - \hat{y}_i)_{w_2}$$

$$= -(y - \hat{y}_i) . \left[\sigma(w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2) + b_3) \right]_{w_2}$$

$$= -(y - \hat{y}_i) \cdot \sigma'(w_3 . \sigma(w_2 . \sigma(w_1 \cdot x_i + b_1) + b_2) + b_3) . \left[w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 \cdot x + b_1) + b_2) + b_3 \right]_{w_2}$$

$$= -(y - \hat{y}_i) \cdot \hat{y}_i \cdot (1 - \hat{y}_i) \cdot w_3 \left[\sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2) \right]_{w_2}$$

$$= -(y - \hat{y}_i) \cdot \hat{y}_i \cdot (1 - \hat{y}_i) \cdot w_3 \cdot \left[\sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2) \cdot (1 - \sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2)) \cdot \sigma(w_1 \cdot x_i + b_1) \right]$$

Figure 8: Derivada w2

$$\frac{\partial \varepsilon}{\partial w_3} = \left[1/2 \, (y - \underbrace{\sigma(w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 * x_i + b_1) + b_2) + b_3))}_{\hat{y}_i} \right]^2 = (y - \hat{y}_i)(y - \hat{y}_i)_{w_3}$$

$$= -(y - \hat{y}_i) \cdot \left[\sigma(w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2) + b_3) \right]_{w_3}$$

$$= -(y - \hat{y}_i) . \hat{y}_i (1 - \hat{y}_i) \cdot \left[w_3 \cdot \sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2) + b_3 \right]_{w_3}$$

$$= -(y - \hat{y}_i) \cdot \hat{y}_i \cdot (1 - \hat{y}_i) \cdot \sigma(w_2 \cdot \sigma(w_1 \cdot x_i + b_1) + b_2)$$

Figure 9: Derivada w3

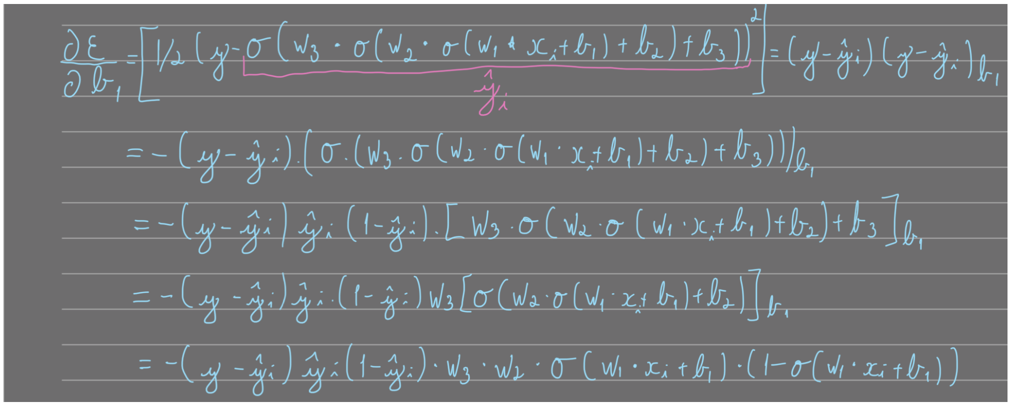

Figure 10: Derivada b1

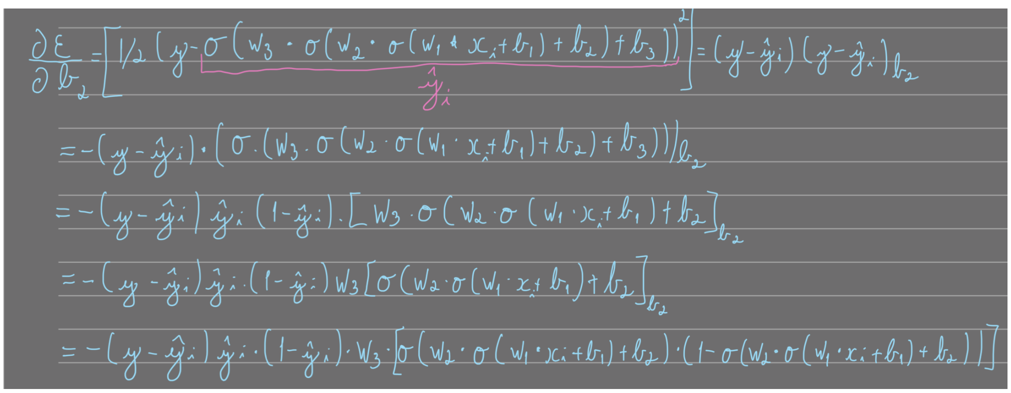

Figure 11: Derivada b2

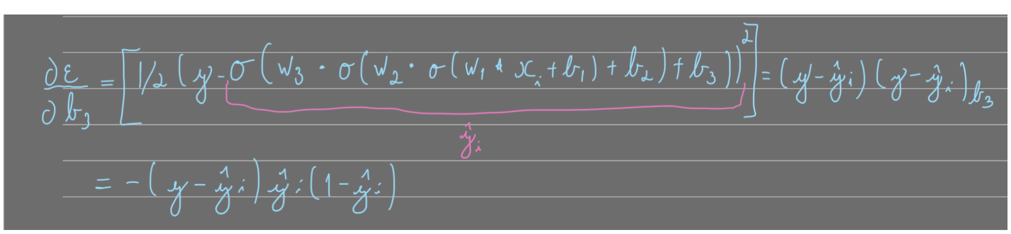

Figure 12: Derivada b3

6 Parâmetros de Teste

Os valores utilizados para w_1, w_2, w_3, b_1, b_2, b_3, taxa de aprendizado (lr) e erro foram:

- $w_1 = 0.05$
- $w_2 = 0.07$
- $w_3 = 0.03$
- $b_1 = -0.02$
- $b_2 = 0.01$
- $b_3 = -0.03$
- $lr = 0.1$
- $err = 0.001$

7 Conclusões

Dos resultados obtidos é possivel ver que o treinamento foi realizado e predição da rede neural atua, com isso classificando os números como primos ou não e pares ou ímpares, como os exemplos abaixo:

```
Número -2900 é primo: False
Número -2900 é par: False
Número 10501 é primo: True
Número 10501 é par: False
Número 41 é primo: True
Número 41 é par: False
Número 76 é primo: False
Número 76 é par: True

Digite um número para classificar ou -1 para sair: 13
Número 13 é primo: True
Número 13 é par: False

Digite um número para classificar ou -1 para sair: 12
Número 12 é primo: False
Número 12 é par: True

Digite um número para classificar ou -1 para sair: 999
Número 999 é primo: True
Número 999 é par: False

Digite um número para classificar ou -1 para sair: 55
Número 55 é primo: True
Número 55 é par: False

Digite um número para classificar ou -1 para sair: 9896
Número 9896 é primo: False
Número 9896 é par: True

Digite um número para classificar ou -1 para sair: -1
```

Figure 13: Exemplos de respostas

8 Link do GitHub

Repositório: *https : //github.com/AugustoRC/Redes_neurais*

References

[1] Iezzi, Gelson. *Fundamentos da Matemática Elementar*, volume 1.