

Victor Rodrigues de Oliveira

**Descobrimdo Redes Neurais Artificiais - Minicurso sobre  
os modelos: Perceptron, Adaline e Perceptron  
Multicamadas**

**Dissertação de Mestrado** apresentada ao  
Instituto Federal de Educação, Ciência e Tec-  
nologia de São Paulo, para obtenção do título  
de Mestre em Matemática.

Instituto Federal de Educação Ciência e Tecnologia de São Paulo

Campus São Paulo

Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT

Orientador: Prof. Dr. Marco Aurélio Granero Santos

São Paulo - SP

2025

# Lista de ilustrações

Figura 1 – Modelo ilustrativo de uma aplicação do Neurônio de McCulloch-Pitts com componentes eletrônicos. . . . .	1
Figura 2 – Plano dividido em duas regiões pela reta de equação $2x - y = 1$ . . . . .	2
Figura 3 – Diagrama do funcionamento do neurônio de McCulloch-Pitts. . . . .	3
Figura 4 – Uma função booleana “E” . . . . .	5
Figura 5 – Diagrama das etapas a serem executadas pelo neurônio artificial. . . . .	5
Figura 6 – Gráficos da função “E” com diferentes limiares. . . . .	7
Figura 7 – Representação de uma função “OU”. . . . .	7
Figura 8 – Gráfico gerado com os pesos sinápticos designados para função “OU”. . . . .	9
Figura 9 – Disposição dos pontos para uma função “XOU” . . . . .	10

# Lista de tabelas

Tabela 1	–	Tabela verdade da função “E”. . . . .	4
Tabela 2	–	Combinações lineares para função “E”. . . . .	6
Tabela 3	–	Tabela verdade da função “OU”. . . . .	8
Tabela 4	–	Combinações lineares dos pesos com as variáveis. . . . .	8



# Sumário

1	NEURÔNIO DE MCCULLOCH E PITTS . . . . .	1
1.1	O funcionamento do Neurônio de McCulloch-Pitts . . . . .	3
1.2	Implementando funções booleanas com o neurônio de McCulloch-Pitts . . . . .	4
	Bibliografia . . . . .	13



# 1 Neurônio de McCulloch e Pitts

Os pesquisadores Warren McCulloch <sup>1</sup> e Walter Pitts <sup>2</sup>, em 1943, apresentaram um modelo teórico que usa álgebra booleana, com “0” e “1” indicando sinais negativos e positivos, respectivamente, para simular o funcionamento de um neurônio biológico e aplicá-lo na tomada de decisões.

Este modelo foi apresentado no artigo “[A logical calculus of the ideas immanent in nervous activity \(1943\)](#)”. Esse trabalho foi um marco, considerado por muitos o inicial, para o desenvolvimento do campo de estudos relativos às Redes Neurais e Inteligência Artificial. O neurônio de McCulloch-Pitts é o que se chama de discriminador binário, ou seja, é um classificador de dados para duas classes distintas (Kovács, 2023).

Com o uso deste neurônio artificial, foram elaboradas máquinas eletrônicas em que o modelo do neurônio era adaptado com uso de componentes eletrônicos como resistores ajustáveis e sensores ópticos, e desempenhava o papel de identificador de padrões, como pode ser visto na Figura 1 (Silva; Spatti; Flauzino, 2016).

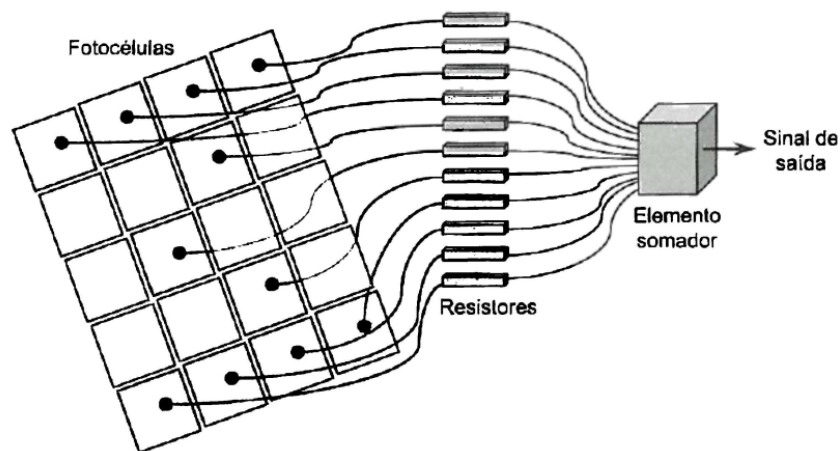


Figura 1 – Modelo ilustrativo de uma aplicação do Neurônio de McCulloch-Pitts com componentes eletrônicos. Fonte: Silva, Spatti e Flauzino (2016, p.58)

Para possibilitar a identificação de padrões que o Neurônio de McCulloch-Pitts é capaz de executar, que é relativa a categorização destes dados em uma de duas classes distintas, é necessário que os dados também sejam separáveis em duas categorias.

<sup>1</sup> Warren Sturgis McCulloch, nascido em 1898 e falecido em 1969, foi um neurofisiologista, era conhecido pelo seu trabalho sobre os fundamentos de certas teorias do cérebro e pela sua contribuição para a área da cibernética (Tappert, 2019).

<sup>2</sup> Walter Pitts, nascido em 1923 e falecido em 1969, foi um matemático formado no MIT. Escreveu diversos artigos sobre métodos de simular, ou até reproduzir, o funcionamento do cérebro humano por meio da linguagem matemática e de máquinas elétricas. Visto em [https://home.csulb.edu/~cwallis/artificialn/walter\\_pitts.html](https://home.csulb.edu/~cwallis/artificialn/walter_pitts.html). Acesso em: 25 de fevereiro de 2025.

Esta divisão é realizada de modo que, sendo possível dispor os dados em um plano cartesiano em duas dimensões, eles seriam completamente separados por uma reta. Por conta desta característica, o Neurônio de McCulloch-Pitts também é chamado de discriminador, ou separador linear.

A Figura 2 apresenta um plano cartesiano no qual uma reta de equação  $2x - y = 1$  é representada, possibilitando uma discussão sobre dos espaços que ela delimita.

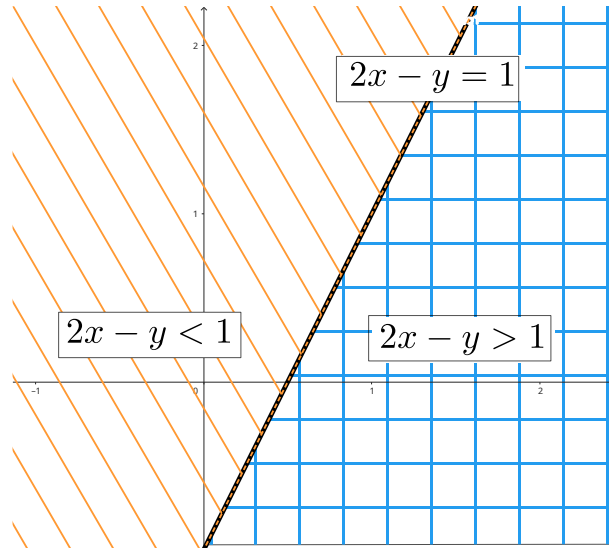


Figura 2 – Plano dividido em duas regiões pela reta de equação  $2x - y = 1$ .  
Elaborado pelo autor.

Analisando o plano no  $\mathbb{R}^2$  exemplificado na Figura 2, a região laranja representa os pares de pontos  $(x, y)$  em que  $2x - y < 1$  e a região azul em que  $2x - y > 1$  enquanto a reta  $2x - y = 1$  é a separadora linear.

De um modo geral, para dados que podem ser dispostos no plano  $\mathbb{R}^2$ , ou seja, com duas variáveis, o discriminador linear é uma reta que faz a separação do plano em outros dois semiplanos com os dados contidos em um ou outro deles.

Com três variáveis, ou seja, no espaço tridimensional  $\mathbb{R}^3$ , uma superfície plana ou simplesmente um plano faz o papel do discriminador linear. Em dimensões maiores, e com mais variáveis, o discriminador linear que separa o espaço  $\mathbb{R}^n$  em duas regiões distintas é comumente chamado de hiperplano.

É importante enfatizar que a separação linear realizada deverá ser apresentada conforme as características relativas aos dados. Eventualmente, pode não ser possível realizar a separação com retas, planos ou hiperplanos, sendo necessária a utilização de outras curvas e superfícies para tal. Porém, este tipo de separação não tem possibilidade de ser executada com somente um neurônio de McCulloch-Pitts (Minsky; Papert, 1988).



## 1.1 O funcionamento do Neurônio de McCulloch-Pitts

A implementação do neurônio artificial com componentes eletrônicos, exemplificado na Figura 1, receberá os dados cuidadosamente tratados para serem lidos de modo eficiente pelos sensores, e então apresentará uma saída, um valor numérico, que estará associado a uma categoria.

O funcionamento do neurônio artificial de McCulloch-Pitts pode ser apresentado como no diagrama da Figura 3.

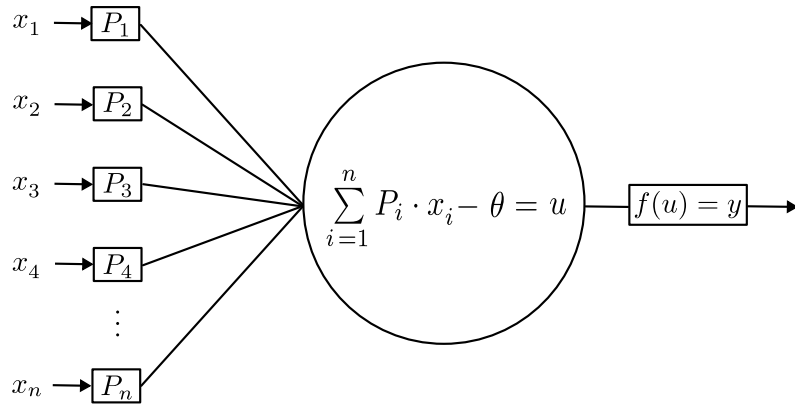


Figura 3 – Diagrama apresentando as entradas e o funcionamento do neurônio artificial de McCulloch-Pitts. Elaborada pelo autor.

As variáveis  $x_1, x_2, \dots, x_n$  representam a informação externa que, quando traduzida para a linguagem matemática, serão repassadas ao neurônio artificial. Estas informações são valores que podem se originar de medições objetivas de fenômenos e características em determinados contextos ou de outras fontes como conjuntos numéricos.

Os valores  $P_1, P_2, \dots, P_n$  são chamados de pesos sinápticos e são ponderadores da importância das variáveis de mesmo índice. Isto significa que os pesos reforçam determinadas características apresentadas pelas variáveis e, atribuem maior ou menor importância para elas com o objetivo de realizar a categorização.

O símbolo  $\sum$  é usado para indicar a soma, neste caso relativa ao produto de cada entrada com seu respectivo peso sináptico, estabelecendo uma combinação linear, com o índice  $i$  variando de 1 até o número  $n$  que é máximo de variáveis fornecidas pelas amostras ao neurônio artificial.

A letra grega  $\theta$  representa um número que comumente é chamado de limiar, *bias* ou *threshold*, que atua como um valor de referência em relação à combinação linear obtida.

A função de ativação  $f$  recebe este valor, já considerando o  $\theta$ , e então devolve uma saída  $y$  que, ao ser analisada pelo operador da Rede Neural Artificial, indicará a atribuição das categorias. A classificação, neste caso, é realizada apresentando duas possibilidades

distintas e a função de ativação que determinará, a partir dos valores reais apresentados por  $u$ , as saídas binárias (Kovács, 2023).

Em resumo, para o neurônio de McCulloch-Pitts, a soma dos produtos dos valores de entradas por seus respectivos pesos deve atingir um valor maior do que  $\theta$  para que a função de ativação possa devolver o valor associado a um conjunto e, quando a soma não superar esse limiar, a função de ativação devolverá o valor associado a outro conjunto. Em se tratando de redes neurais envolvendo o neurônio de McCulloch-Pitts o limiar também será o parâmetro de referência na função de ativação. Os pesos sinápticos e limiar tem seus valores fornecidos durante a implementação da rede neural (Kovács, 2023).

## 1.2 Implementando funções booleanas com o neurônio de McCulloch-Pitts

Como forma de exemplificar o funcionamento do neurônio artificial, serão implementadas as funções booleanas “E” e “OU” com duas variáveis. Estas duas funções podem ser implementadas com o uso de somente um neurônio artificial. Ambas são caracterizadas por receberem em suas entradas dois valores binários “0” e “1” e devolverem uma saída binária que também será um dentre estes dois valores.

Considerando que estes dados estão com duas variáveis, é possível realizar sua disposição em um plano cartesiano com eixos  $x_1$  e  $x_2$ , de forma que eles sejam considerados pares ordenados. Seguindo nesta adaptação, pode-se então assumir que categorizar estes dados linearmente é equivalente a obter uma reta que divide o plano em duas regiões de forma que em cada uma delas estejam os pontos que apresentam a mesma saída pela função de ativação, como pode ser observado na Figura 4 ao considerar que os pesos sinápticos do neurônio podem ser interpretados, em um caso com duas variáveis, como coeficientes da equação de uma reta.

A função booleana “E” apresentará o valor “1” somente quando nas duas entradas este valor estiver presente. Considerando  $f$  como a função “E”, a tabela verdade para as entradas  $x_1$  e  $x_2$  assumindo valores binários é dada por:

$x_1$	$x_2$	$f(x_1, x_2) = x_1 \wedge x_2$
0	0	0
0	1	0
1	0	0
1	1	1

Tabela 1 – Tabela verdade de função booleana “E”.

onde o sinal “ $\wedge$ ” indica a atuação da função booleana “E” nas variáveis  $x_1$  e  $x_2$ .

Dispondo os pontos referentes às entradas da função “E” no plano cartesiano, tal que os pares de entrada que têm saída “1” são representados por um ponto e, os pares de entrada com saída “0” são representados por um círculo, tem-se a representação gráfica do problema associado à função booleana “E” ilustrado na Figura 4.

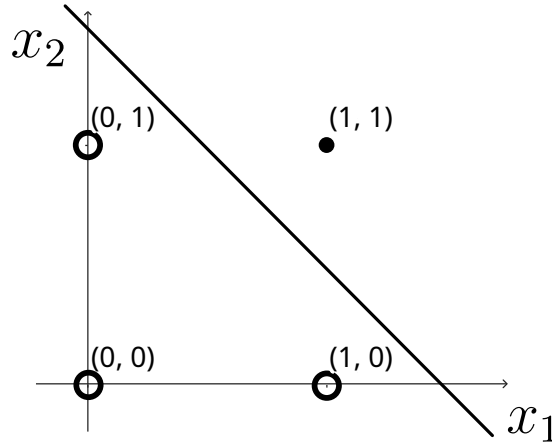


Figura 4 – Representação de uma função booleana “E” dentre as várias possíveis.  
Elaborado pelo autor.

Ainda na Figura 4, a reta apresentada é uma candidata a separadora linear enquanto os pontos  $(0,0)$ ,  $(0,1)$  e  $(1,0)$  estão agrupados em uma região, e o ponto  $(1,1)$  está em outra. A tarefa da Rede Neural Artificial será, então, a de encontrar os coeficientes que esta reta deve ter para que exista uma separação, linear, como a apresentada na Figura 4.

Deste modo, o neurônio de McCulloch-Pitts pode ser representado, para este problema, por meio do diagrama de blocos ilustrado na Figura 5, em que os parâmetros pesos  $P_1$  e  $P_2$  e o limiar  $\theta$  devem ser ajustados pela rede neural de modo a produzir a classificação desejada.

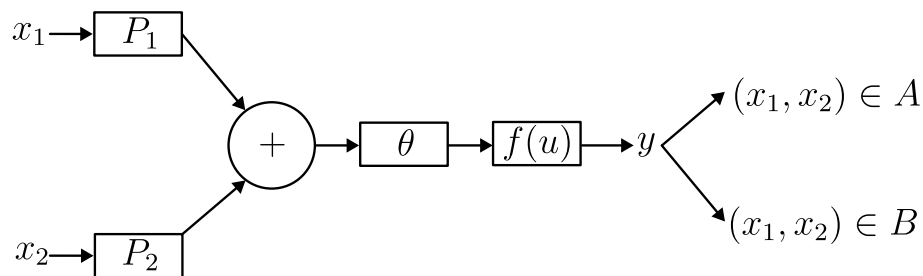


Figura 5 – Diagrama das etapas a serem executadas pelo neurônio artificial.  
Elaborado pelo autor.

É possível escrever a equação (1.1) com pesos sinápticos ponderando as variáveis das amostras das entradas  $x_1$  e  $x_2$  utilizando a Tabela 1.

$$\begin{cases} P_1 \cdot 0 + P_2 \cdot 0 = 0 \\ P_1 \cdot 0 + P_2 \cdot 1 = 0 \\ P_1 \cdot 1 + P_2 \cdot 0 = 0 \\ P_1 \cdot 1 + P_2 \cdot 1 = 1 \end{cases} \quad (1.1)$$

A Equação 1.1 apresenta um sistema de equações lineares. Este sistema é chamado de **Sobredeterminado** e, geralmente, não apresenta valores que satisfaçam todas as equações simultaneamente devido ao número de equações, também chamado de restrições, ser maior que o número de variáveis (Anton; Rorres, 2012, p.241).

Adotando, por exemplo,  $P_1 = P_2 = 1$  o sistema linear da Equação (1.1), considerando também o limiar  $\theta$  resulta em:

$(x_1, x_2)$	$P_1 \cdot x_1 + P_2 \cdot x_2 - \theta = u$
(0, 0)	$1 \cdot 0 + 1 \cdot 0 - \theta = -\theta$
(0, 1)	$1 \cdot 0 + 1 \cdot 1 - \theta = 1 - \theta$
(1, 0)	$1 \cdot 1 + 1 \cdot 0 - \theta = 1 - \theta$
(1, 1)	$1 \cdot 1 + 1 \cdot 1 - \theta = 2 - \theta$

Tabela 2 – Combinação linear entre os pesos e variáveis dos dados da função “E”.  
Elaborado pelo autor.

O limiar  $\theta$  opera como um critério para a rede e deve ser escolhido (ou ajustado) de forma que as saídas apresentem enquadramento a uma ou outra categoria. Desse modo, pode-se adotar um valor para  $\theta$  no intervalo  $]1, 2]$  pois, de acordo com a Tabela 2, a combinação linear que deverá estar associada à saída “1” é somente a da última linha enquanto todos os valores das outras linhas devem ser associados a “0”.

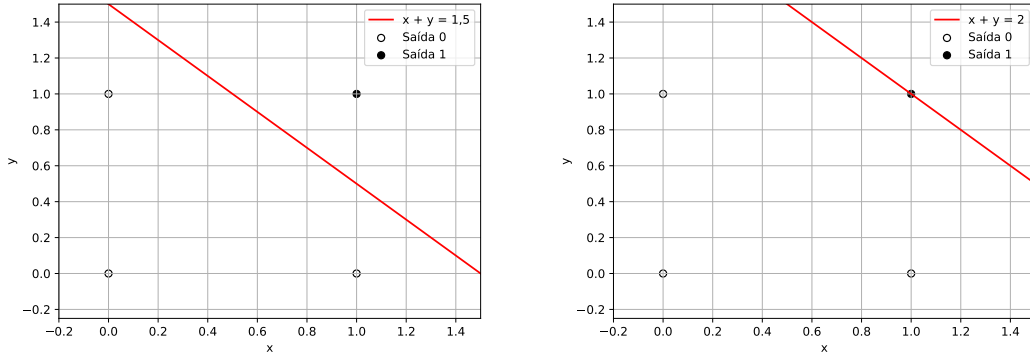
Adotando-se então o limiar  $\theta = 1,5$ , a função de ativação pode ser escrita como:

$$y = f(u) = \begin{cases} 1 & \text{se } u \geq 0 \\ 0 & \text{se } u < 0. \end{cases} \quad (1.2)$$

Sendo assim, os coeficientes da reta de equação:  $ax + by = c$  para que ela atue como separador linear para o conjunto de quatro pontos com correspondência com a função booleana “E” podem ser:  $a = b = 1$  e o coeficiente independente  $c = 1,5$ .

Este resultado pode ser observado na Figura 6a, em que a reta  $x + y = 1,5$  divide o plano em duas regiões: uma delas,  $x + y < 1,5$ , contendo os pontos (0, 0), (0, 1) e (1, 0) cujas saídas da rede neural são “0”; e, outra região,  $x + y \geq 1,5$ , contendo o ponto (1, 1) cuja saída é “1”.

Explorando um pouco mais esta situação, caso a reta tenha coeficiente independente  $c$  igual a 2, a expressão para função de ativação apresentada na Equação (1.2) ainda permaneceria válida, porém, o ponto  $(1, 1)$  estaria contido na reta de equação  $x + y = 2$ , como visto na Figura 6b.



(a) Gráfico da função “E” representado pela reta  $x + y = 1,5$ .  
 (b) Gráfico da função “E” representado pela reta  $x + y = 2$ .

Figura 6 – Gráficos da função “E” com diferentes limiares. Elaborado pelo autor.

Para representar a função booleana “OU”, pode-se adotar a mesma estratégia utilizada no exemplo referente a função booleana “E”, ou seja, obter um separador linear conforme ilustrado na Figura 7. Neste caso, também pode-se utilizar os valores recebidos dos pesos sinápticos do neurônio como coeficientes da equação da reta separatriz.

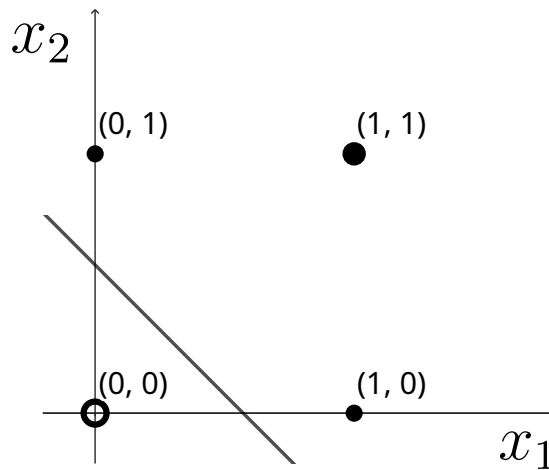


Figura 7 – Representação gráfica de uma função booleana “OU”. Elaborado pelo autor

A tabela verdade para as entradas  $x_1$  e  $x_2$  assumindo valores binários é observada na Tabela 3 em que o sinal de  $\vee$  está indicando a atuação da função booleana “OU” nas variáveis  $x_1$  e  $x_2$ .

$x_1$	$x_2$	$f(x_1, x_2) = x_1 \vee x_2$
0	0	0
0	1	1
1	0	1
1	1	1

Tabela 3 – Tabela verdade da função booleana “OU”.

Escrevendo as combinações lineares entre as entradas que  $x_1$  e  $x_2$  podem assumir e os pesos  $P_1$  e  $P_2$ , tem-se

$$\begin{cases} P_1 \cdot 0 + P_2 \cdot 0 = 0 \\ P_1 \cdot 0 + P_2 \cdot 1 = 1 \\ P_1 \cdot 1 + P_2 \cdot 0 = 1 \\ P_1 \cdot 1 + P_2 \cdot 1 = 1 \end{cases} \quad (1.3)$$

A Equação (1.3) também é um sistema linear sobredeterminado. Similar ao que foi feito no exemplo da função booleana “E”, também pode-se estabelecer os pesos sinápticos convenientemente. Inclusive, seria possível repetir aqui os pesos sinápticos adotados na Tabela 2.

Porém, para evidenciar as diferentes possibilidades de pesos que podem ser utilizados, e a conveniência do limiar escolhido, serão adotados dois valores iguais a  $\frac{1}{2}$ , ou seja,  $P_1 = P_2 = \frac{1}{2}$ . Realizando os procedimentos, linha por linha, tem-se:

$(x_1, x_2)$	$P_1.x_1 + P_2.x_2 - \theta = u$
(0, 0)	$\frac{1}{2}.0 + \frac{1}{2}.0 - \theta = -\theta$
(0, 1)	$\frac{1}{2}.0 + \frac{1}{2}.1 - \theta = \frac{1}{2} - \theta$
(1, 0)	$\frac{1}{2}.1 + \frac{1}{2}.0 - \theta = \frac{1}{2} - \theta$
(1, 1)	$\frac{1}{2}.1 + \frac{1}{2}.1 - \theta = 1 - \theta$

Tabela 4 – Combinação linear dos pesos com as variáveis.

Desse modo, percebe-se que o limiar  $\theta$  pode ser um valor entre 0 e  $\frac{1}{2}$ , pois, de acordo com a Tabela 4, a combinação linear que deverá ser associada a “0” é somente a da primeira linha, enquanto que todos os valores das outras linhas devem ser associados a “1”.

Adotando  $\theta = \frac{1}{2}$ , a função de ativação pode ser escrita como:

$$y = f(u) = \begin{cases} 1 & \text{se } u \geq 0 \\ 0 & \text{se } u < 0. \end{cases} \quad (1.4)$$

que representa a implementação da função lógica “OU” com um neurônio de McCulloch-Pitts.

A Figura 8 ilustra graficamente os dados de entrada, as saídas e o separador linear dado pela reta de equação:  $\frac{x}{2} + \frac{y}{2} = \frac{1}{2}$ .

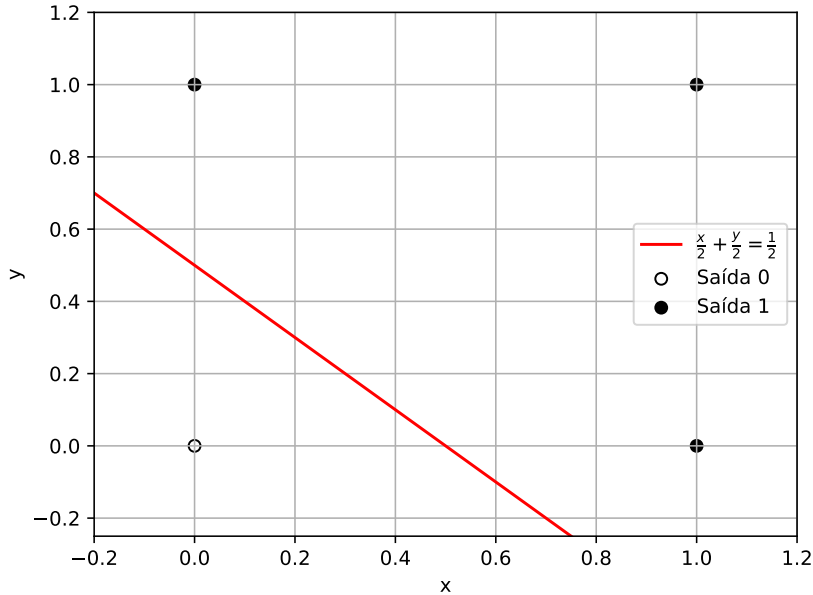


Figura 8 – Gráfico da função “OU” representado pela reta de equação  $\frac{x}{2} + \frac{y}{2} = \frac{1}{2}$ .  
Elaborado pelo autor.

Os pesos sinápticos  $P_1$  e  $P_2$  de ambos neurônios foram atribuídos arbitrariamente, e os limiares  $\theta$ , de modo conveniente. A observação do atendimento das condições da função booleana estabelece alguns limites para esta escolha, mas, teoricamente, existem infinitas combinações de pesos e limiares, e consequentemente, de coeficientes da reta que realizem a separação do plano do modo necessário a representar corretamente cada uma das funções exemplificadas.

É importante notar que a função de ativação utilizada nas duas implementações esteve diretamente ligada com o limiar  $\theta$  de maneira que este valor foi explicitamente adotado no critério desta função. Em ambos os casos, pode-se identificar o uso da função sinal, ou degrau,<sup>3</sup> que pode ser expressa, para algum número  $u$  real,

$$H(u) = \begin{cases} 1 & \text{se } u \geq \theta \\ 0 & \text{se } u < \theta \end{cases}, \quad (1.5)$$

o que resulta nas saídas da função de ativação serem somente valores binários.

Nos exemplos apresentados anteriormente foi possível realizar a separação dos dados com somente um neurônio pois as variáveis nas entradas eram binárias e linearmente separáveis, ou seja, por meio de combinações lineares dos pesos com as entradas, são capazes serem classificados em uma ou outra categoria.

<sup>3</sup> Também conhecida como função de Heaviside, em homenagem ao matemático e engenheiro eletricista Oliver Heaviside (1850 - 1925)

Nem todos os conjuntos de dados são linearmente separáveis e isso faz com que um neurônio de McCulloch-Pitts possa não apresentar uma classificação adequada.

Para duas entradas,  $x_1$  e  $x_2$ , a função booleana *ou-exclusivo*, abreviada como “XOU”, apresenta valores binários “1” para combinações que possuam exclusivamente uma ou outra entrada igual à “1”, e “0” caso contrário, conforme ilustrado na Figura 9.

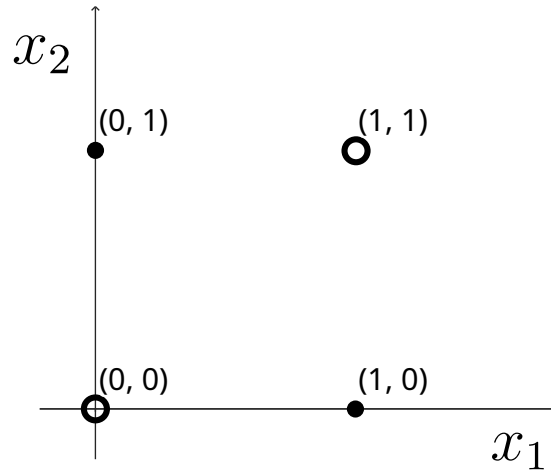


Figura 9 – Disposição dos pontos no plano com saídas associadas à função booleana “XOU”. Elaborado pelo autor.

Esta função não pode ser implementada com somente um neurônio artificial (Minsky; Papert, 1988). Ainda segundo Widrow e Lehr (1990) e Kovács (2023), na impossibilidade de um neurônio representar uma função booleana, uma combinação de funções, e consequentemente de neurônios, será capaz de realizá-la. Entretanto, este tipo de rede consiste no uso de mais de uma camada de neurônios, assunto tratado no capítulo ??.

Um fator importantíssimo para o neurônio de McCulloch-Pitts assim apresentado, é a atribuição dos pesos sinápticos e limiares manualmente. Nos exemplos que foram utilizados isto não demandou muita criatividade, porém, este neurônio artificial é capaz de realizar a categorização de conjuntos de dados linearmente separáveis com centenas, ou milhares de dados.

Dessa forma, nestes casos há a necessidade da atribuição automatizada, por meio de algoritmos, de pesos sinápticos adequados. Isto é chamado de treinamento da Rede Neural. Nos capítulos ?? e ??, serão apresentados dois modos distintos de se realizar o treinamento e a obtenção destes pesos sinápticos de modo automático.







# Bibliografia

ANTON, Howard; RORRES, Chris. **Álgebra Linear com aplicações**. 10. ed. Porto Alegre: Bookman, 2012. P. 576.

KOVÁCS, Zsolt László. **Redes Neurais Artificiais: Fundamentos e Aplicações**. 4. ed.: Livraria da Física, 2023. P. 176.

MCCULLOCH, Warren S.; PITTS, Walter. **A logical calculus of the ideas immanent in nervous activity**. *The Bulletin of Mathematical Biophysics*, Springer Science e Business Media LLC, v. 5, n. 4, p. 115–133, dez. 1943. ISSN 1522-9602. DOI: [10.1007/bf02478259](https://doi.org/10.1007/bf02478259).

MINSKY, Marvin; PAPERT, Seymour Aubrey. **Perceptrons: An Introduction to Computational Geometry**. Expanded: The MIT Press, 1988.

SILVA, Ivan Nunes da; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. **Redes Neurais Artificiais Para Engenharia e Ciências Aplicadas: Fundamentos Teóricos e Aspectos Práticos**. 2. ed. São Paulo: Artliber, 2016. P. 431.

TAPPERT, Charles C. **Who Is the Father of Deep Learning?** *International Conference on Computational Science and Computational Intelligence*, 2019.

WIDROW, Bernard; LEHR, Michael A. **30 years of adaptive neural networks: perceptron, Madaline, and backpropagation**. *Proceedings of the IEEE*, v. 78, n. 9, p. 1415–1442, 1990. DOI: [10.1109/5.58323](https://doi.org/10.1109/5.58323).