

# A MATEMÁTICA POR TRÁS DAS REDES NEURAIS

BARONI, Itanimá<sup>1</sup>

PADILHA, Eliandro José

## RESUMO

Dentre os diversos algoritmos de Aprendizado de Máquina (AM), as Redes Neurais Artificiais - RNA têm um destaque especial em virtude de serem modelos baseados no comportamento do cérebro humano. Elas foram concebidas para se assemelharem ao comportamento dos neurônios humanos, no que tange à obtenção do conhecimento através de um processo de aprendizado, e especialmente na utilização de conexões neurais, conhecidas como sinapses. Este artigo busca ilustrar a matemática por trás das redes neurais, de modo exemplificativo, por intermédio de dois exemplos bem simplificados, mas que demonstram a importância do saber matemático, seja na seleção dos melhores algoritmos para a solução dos problemas, ou na otimização do algoritmo escolhido, tentando desmistificar a mera utilização dos *frameworks* disponíveis atualmente. A metodologia utilizada nesse trabalho foi o levantamento bibliográfico e documental sobre o objeto de estudo a fim de trazer uma visão ampla a respeito dos conceitos matemáticos das Redes Neurais Artificiais, e a rede neural utilizada será a rede Perceptron, escolhida pela sua simplicidade, limitando à sua estrutura ao número mínimo de neurônios, possibilitando abstraímos da complexidade dos algoritmos que compõe as redes neurais para nos atermos ao entendimento matemático.

**Palavras-chave:** Redes Neurais. Matemática. Algoritmo.

## 1 INTRODUÇÃO

As máquinas sempre nos ajudaram desde o advento da Revolução Industrial. Elas não apenas aumentam nossa produtividade, mas também constituem um fator fundamental na determinação da economia de uma nação. Elas passaram por uma

---

<sup>1</sup> Aluno do Curso de Bacharelado em Matemática do Centro Universitário Internacional UNINTER. Artigo apresentado como Trabalho de Conclusão de Curso - 07/2021 – RU 2607819.

série de melhorias nos últimos séculos e transformaram-se em uma ampla variedade desde sua primeira aparição. Mas uma coisa permaneceu mais ou menos igual, que é a sua dependência explícita aos comandos humanos.

Nos últimos anos, pesquisas estão sendo realizadas para infligir inteligência às máquinas. Assim, os termos “Inteligência Artificial” e “Aprendizado de Máquina” passaram a existir.

A definição de Inteligência Artificial (IA) é causa de muitos debates entres os pesquisadores do assunto, não por causa da definição do que é “Artificial”, mas sim do que é considerado “Inteligência”, haja vista ser este um conceito subjetivo.

Neste contexto, diante de tantas divergências Norvig e Russel (2013), propuseram a separação das definições em duas dimensões. Na parte superior do quando abaixo estão as que se relacionam a processos de pensamento e raciocínio, e as definições da parte inferior se referem ao comportamento. As definições do lado esquerdo tratam da fidelidade ao desempenho humano, e as do lado direito tratam do conceito ideal de inteligência, chamado de racionalidade.

Quadro 1 – Definições de IA

<b>Pensando como um humano</b>	<b>Pensando racionalmente</b>
“O novo e interessante esforço para fazer os computadores pensarem (...) máquinas com mentes, no sentido total e literal.” (Haugeland, 1985)	“O estudo das faculdades mentais pelo uso de modelos computacionais.” (Charniak e McDermott, 1985)
“[Automatização de] atividades que associamos ao pensamento humano, atividades como a tomada de decisões, a resolução de problemas, o aprendizado...” (Bellman, 1978)	“O estudo das computações que tornam possível perceber, raciocinar e agir.” (Winston, 1992)
<b>Agindo como seres humanos</b>	<b>Agindo racionalmente</b>
“A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.” (Kurzweil, 1990)	“Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (Poole et al., 1998)
“O estudo de como os computadores podem fazer tarefas que hoje são melhor desempenhadas pelas pessoas.” (Rich and Knight, 1991)	“AI... está relacionada a um desempenho inteligente de artefatos.” (Nilsson, 1998)

Fonte: NORVIG e RUSSEL (2013, p. 25)

A expectativa criada em torno da solução de problemas rotineiros, além da possibilidade real de melhorias de desempenho nas tarefas complexas fez com que houvesse, a partir da década de 1970, uma disseminação das técnicas baseadas em Inteligência Artificial, e uma década depois Rowe (1988) previu o que deveríamos esperar.

A inteligência artificial inclui:

- Fazer com que os computadores se comuniquem conosco em línguas humanas como o inglês, seja por imprimir em um terminal de computador, entender as coisas que digitamos em um terminal de computador, geração de fala, ou compreensão de nossa fala (linguagem natural);
- Fazer com que os computadores SE lembrem de fatos complicados e inter-relacionados e tirem conclusões a partir deles (inferência);
- Obter computadores para planejar seqüências de ações para cumprir metas (planejamento);
- Obter computadores para nos oferecer conselhos com base em regras complicadas para várias situações (sistemas especializados);
- Fazer com que os computadores olhem pelas câmeras e vejam o que está lá (visão);
- Fazer com que os computadores se movam e movam objetos no mundo real (robótica). ROWE (1988).

Com o passar do tempo, houve um aumento na complexidade dos problemas a serem resolvidos computacionalmente, além do aumento exponencial do volume de dados gerados nos mais diversos setores da economia, o que tornou imperioso a criação de ferramentas computacionais capazes de resolver os problemas de forma autônoma, com o mínimo de intervenção humana. E para que isso ocorresse essas ferramentas deveriam ter a habilidade de, a partir de experiências passadas, criarem autonomamente uma hipótese ou função que fosse capaz de resolver o problema.

Ao processo computacional de aprendizado com experiências passadas, por meio de um conjunto de exemplos deu-se o nome de Aprendizado de Máquina (AM). Os algoritmos que compõe esta ferramenta são capazes de aprender a partir de um conjunto de dados, e propor uma solução para o problema. Segundo Mitchell (1997, p. 2) Aprendizado de Máquina (AM) é “A capacidade de melhorar o desempenho na realização de alguma tarefa por meio da experiência”.

**Definição:** Diz-se que um programa de computador **aprende** com a experiência E com respeito a alguma classe de tarefas T e medida de desempenho P, se o seu desempenho nas tarefas em T, conforme medido por P, melhora com a experiência E. MITCHELL (1997, p. 2)

As Redes Neurais são um dos inúmeros algoritmos disponíveis de Aprendizado de Máquina. Dentre as suas características mais marcantes estão sua estruturação semelhante à rede de neurônios humanos, e a sua capacidade de aprendizado, permitindo a tomada de decisões com base no que aprendeu.

O presente trabalho pretende abordar o que dá “vida” às redes neurais, permitindo avaliar se atualmente, tendo à nossa disposição inúmeras bibliotecas e

*frameworks* especializados de alto nível, precisaríamos entender a matemática por trás das mesmas.

## 2. REDES NEURAIS

Um dos algoritmos de Aprendizado de Máquina que vêm apresentado grande destaque nos últimos anos são as Redes Neurais Artificiais (RNA). O sucesso se deve ao fato de que este algoritmo possibilita a resolução de diversos problemas complexos, nas mais diversas situações da vida real.

Uma característica bastante explorada é a capacidade de aprendizado das redes neurais, permitindo a modelagem de relações entre os dados de entrada e saída, independente do tipo que estes se apresentem, ou seja, podem ser dados não-lineares e complexos, estruturados ou não, imagens, além de séries temporais.

O conceito de "Rede Neural" faz analogia a um algoritmo de aprendizado profundo estruturado de forma semelhante à organização dos neurônios no cérebro humano. Haykin (2001) apresentou a seguinte definição de uma rede neural sob a ótica de uma máquina adaptativa:

Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso. Ela se assemelha ao cérebro em dois aspectos:

1. O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem.
2. Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido. HAYKIN (2001, p. 28).

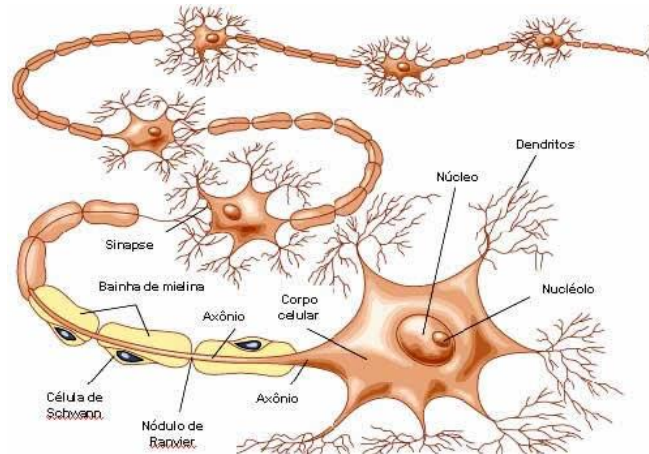
Ludwig e Costa (2007) relacionaram os principais componentes dos neurônios biológicos, de modo a facilitar o entendimento dos seus correlatos nas Redes Neurais Artificiais.

Os principais componentes dos neurônios biológicos são:

- Os dendritos, que têm por função receber os estímulos transmitidos pelos outros neurônios;
- O corpo do neurônio, também chamado de soma, que é responsável por coletar e combinar informações vindas de outrem neurônios;

- O axônio, que é constituído de uma fibra tubular que pode alcançar até alguns metros e é responsável por transmitir os estímulos para outras células. LUDWIG e COSTA (2007, p. 10)

Figura 1: Rede de Neurônios



Fonte: "Células nervosas" em Só Biologia. Virtuuous Tecnologia da Informação, 2008-2021.

Um neurônio biológico recebe sinais por intermédio dos diversos dendritos, que serão examinados e encaminhados para o axônio, e poderão ser encaminhados para o próximo neurônio, ou não (*threshold*). Nesta passagem pelo neurônio, o sinal poderá ter sua intensidade aumentada ou diminuída, dependendo do dendrito que o encaminhou, pois existe um peso associado a cada condutor, que será utilizado para multiplicar o sinal. As memórias são os pesos.

De acordo com Laudon & Laudon (1996), cada região do cérebro é especializado em uma determinada função, que apresentam características distintas, como arquitetura da rede, número de neurônios, número de sinapses por neurônio, valor dos *thresholds* e dos pesos.

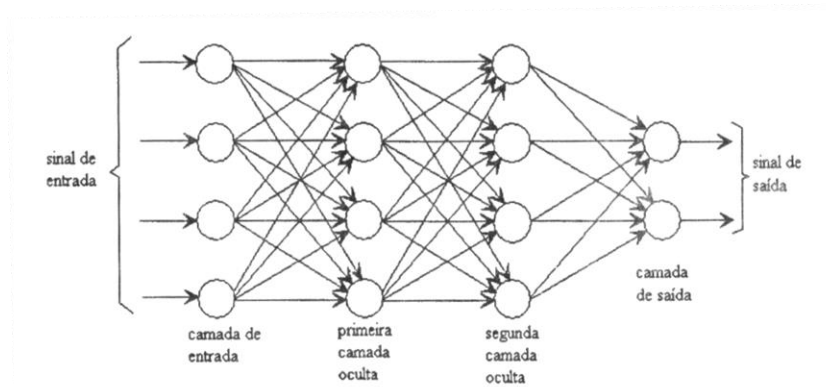
[...] cada região do cérebro é especializada em uma dada função, como processamento de sinais auditivos, sonoros, elaboração de pensamentos, desejos etc. Esse processamento se dá através de redes particulares. Cada região do cérebro possui uma arquitetura de rede diferente: varia o número de neurônios, de sinapses por neurônio, valor dos *thresholds* e dos pesos. LAUDON & LAUDON (1996)

O cérebro no decorrer da sua vida precisa de treinamento para cada tarefa nova a ser executada, de forma a se especializar e aperfeiçoar seu desempenho, e é nesse momento que os pesos estarão sendo definidos. A essa etapa chamou-se memorização.

As Redes Neurais Artificiais são compostas por camadas de nós. Cada nó é projetado para se comportar de forma semelhante a um neurônio no cérebro, sendo a

primeira camada chamada de camada de entrada, seguida por camadas ocultas e, finalmente, a camada de saída. Cada nó na rede neural realiza algum tipo de cálculo, que é passado para outros nós mais profundos na rede neural. A figura abaixo é uma visualização simplificada para demonstrar como isso funciona:

Figura 2: Exemplo de RNA defendida por Ludwig e Costa (2007)



Fonte: LUDWIG e COSTA (2007, p. 15)

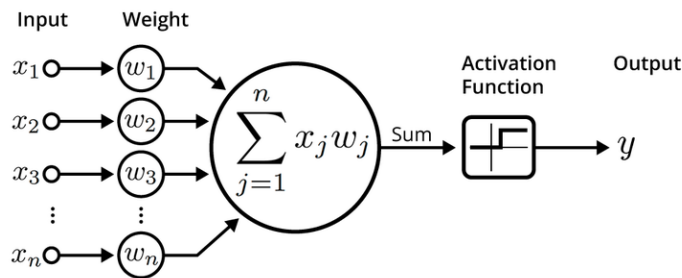
Os neurônios em aprendizagem profunda recebem um ou mais sinais de entrada. Esses sinais podem vir do conjunto de dados brutos ou de neurônios posicionados em uma camada anterior da rede neural, então efetuam alguns cálculos, e enviam alguns sinais de saída para neurônios mais profundos na rede neural por meio da sinapse.

Os neurônios neste tipo de modelo são capazes de ter sinapses que se conectam a mais de um neurônio na camada anterior. Cada sinapse tem um peso associado, que impacta a importância do neurônio precedente na rede neural geral.

Os pesos são um tópico muito importante no campo do aprendizado profundo porque o ajuste dos pesos de um modelo é a principal maneira de treinar os modelos.

Depois que um neurônio recebe suas entradas dos neurônios da camada anterior, ele soma cada sinal multiplicado por seu peso correspondente e os passa para uma função de ativação.

Figura 3: Função de ativação de uma Rede Neural Artificial



Fonte: SHUBH SAXENA (2017).

A função de ativação calcula o valor de saída do neurônio, e esse valor de saída é então passado para a próxima camada da rede neural por meio de outra sinapse – *Função de Transferência*.

### 3. METODOLOGIA

Conceitualmente, metodologia é definida como parte da ciência que se dedica aos procedimentos organizados, e aos métodos empregados pela própria ciência, e de acordo com Fonseca (2002), “A Metodologia é a explicação detalhada de toda ação a desenvolver durante o trabalho de pesquisa”.

Iniciamos o trabalho realizando uma pesquisa abrangente da literatura, com o objetivo de conhecer com maior profundidade o assunto, de modo a torná-lo mais claro, e ser possível abordar questões importantes para a condução da pesquisa. Conforme Fonseca (2002, p. 20) a pesquisa quantitativa é definida como:

A pesquisa quantitativa se centra na objetividade. Influenciada pelo positivismo, considera que a realidade só pode ser compreendida com base na análise de dados brutos, recolhidos com o auxílio de instrumentos padronizados e neutros. A pesquisa quantitativa recorre à linguagem matemática para descrever as causas de um fenômeno, as relações entre variáveis, etc. A utilização conjunta da pesquisa qualitativa e quantitativa permite recolher mais informações do que se poderia conseguir isoladamente. FONSECA (2002, p. 20)

A pesquisa aplicada, quanto à sua natureza, busca produzir conhecimento para aplicações práticas, objetivando solucionar problemas específicos, e com esse propósito, a base inicial da pesquisa foi o levantamento bibliográfico e documental, a fim de termos uma visão mais ampla a respeito dos conceitos matemáticos das Redes Neurais Artificiais, assim como buscar fundamentação teórica para o trabalho.

[...]

A pesquisa bibliográfica é feita a partir do levantamento de referências teóricas já analisadas, e publicadas por meios escritos e eletrônicos, como livros, artigos científicos, páginas de web sites. Qualquer trabalho científico inicia-se com uma pesquisa bibliográfica, que permite ao pesquisador conhecer o que já se estudou sobre o assunto. Existem porém pesquisas científicas que se baseiam unicamente na pesquisa bibliográfica, procurando referências teóricas publicadas com o objetivo de recolher informações ou conhecimentos prévios sobre o problema a respeito do qual se procura a resposta (FONSECA, 2002, p. 31).

[...]

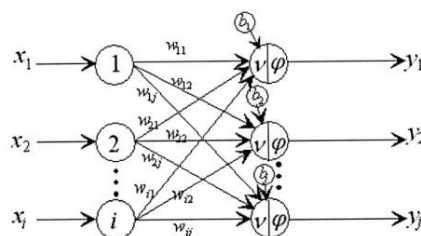
A pesquisa documental trilha os mesmos caminhos da pesquisa bibliográfica, não sendo fácil por vezes distingui-las. A pesquisa bibliográfica utiliza fontes constituídas por material já elaborado, constituído basicamente por livros e artigos científicos localizados em bibliotecas. A pesquisa documental recorre a fontes mais diversificadas e dispersas, sem tratamento analítico, tais como: tabelas estatísticas, jornais, revistas, relatórios, documentos oficiais, cartas, filmes, fotografias, pinturas, tapeçarias, relatórios de empresas, vídeos de programas de televisão, etc. (FONSECA, 2002, p. 32).

Para o desenvolvimento da solução utilizaremos a Rede Neural Perceptron, com apenas uma camada de entrada e uma camada de saída, de modo a exemplificar o funcionamento na prática de uma rede neural, e poderemos nos ater à matemática por trás da mesma.

Como citado anteriormente, os elementos da camada de entrada fazem a distribuição de cada sinal de entrada para todos os neurônios da camada de saída, que é formada por um número de neurônios igual ao número de sinais de saída.

Os neurônios são constituídos por uma função de ativação e uma função de transferência. A função de ativação faz a soma ponderada dos sinais de entrada e a função de transferência determina a saída do neurônio, em função da soma ponderada.

Figura 4: Rede Neural Perceptron



Fonte: LUDWIG e COSTA (2007, p. 28)

Os valores  $w$  representam os pesos sinápticos das ligações entre os neurônios do elemento  $i$  da camada de entrada com o neurônio  $j$  da camada de saída, e simula a eficácia da sinapse entre os neurônios naturais. A função de ativação está representada pela letra  $v$  e tem o formato da equação  $v_j = \sum_{i=1}^n w_{ij}x_i$ .



Neste exemplo a função de transferência será representada pelo símbolo  $\varphi$  e, será utilizada a função degrau, onde  $\varphi$  pode assumir dois valores distintos, conforme abaixo:

$$\varphi(v) = \begin{cases} 1, \forall v \geq 0 \\ 0, \forall v < 0 \end{cases}$$

### 3.1 O ALGORITMO DE APRENDIZAGEM DA REDE NEURAL PERCEPTRON

O ajuste dos pesos sinápticos de uma rede neural permite que, dado um conjunto de sinais de entrada, depois de processá-los através de seus neurônios, apresente o conjunto de sinais de saída desejado, com um nível de erro aceitável.

Este ajuste é obtido, iterativamente, através de um algoritmo de aprendizado, que seguirá os seguintes passos: na partida serão atribuídos aos pesos valores aleatórios e, com eles, apresenta-se um conjunto de sinais de entrada e calcula-se a resposta da rede neural. Logo em seguida, comparam-se os valores calculados com os valores desejados (treinamento supervisionado). Caso o erro não seja aceitável, faz-se o ajuste dos pesos proporcionalmente ao erro e ao valor do sinal de entrada correspondente.

Sabe-se que quanto maior for o erro, maior deverá ser a correção para os pesos. Assim como, quanto maior for o valor do sinal de entrada correspondente a um peso, maior será a influência deste peso na composição dos valores de saída e, por conseguinte, no erro global apresentado, devendo receber uma correção maior.

A expressão que representa todo este processo é a seguinte:

$$w(i,j)_{T+1} = w(i,j)_T + \eta E(j)_T x(i), \text{ onde:}$$

- $w(i,j)_{T+1}$  = valor do peso corrigido;
- $w(i,j)_T$  = valor do peso na iteração anterior;
- $E(j)_T$  = Valor do erro para o neurônio  $j$ ;
- $i$  = índice do sinal de entrada;
- $T$  = iteração;
- $j$  = índice do neurônio;

- $\eta$  = taxa de aprendizado;
- $x(i)$  = sinal de entrada.

O erro  $E(j)$  é a diferença entre o sinal de saída desejado para o neurônio  $j = d(j)$ , e o sinal de saída calculado pela rede neural para aquele neurônio –  $y(j)$ , então  $E(j) = d(j) - y(j)$ .

O erro médio para todos os neurônios da camada de saída na iteração será:  $\varepsilon(T) = \frac{\sum_{j=1}^n |E(j)|}{n}$ , onde  $n$  é o número de neurônios da camada de saída.

O erro médio para todo o conjunto de treinamento será:  $\varepsilon_{med} = \frac{\sum_{T=1}^n \varepsilon(T)}{n}$ , e este valor pode ser utilizado como referência para o encerramento da seção de treinamento pela avaliação do nível de precisão da rede neural.

### 3.1.1 PRIMEIRA APLICAÇÃO

Nesta primeira aplicação da metodologia, a rede neural Perceptron terá dois sinais de entrada e um neurônio na camada de saída. Esta rede neural, após o treinamento, será capaz de classificar quatro indivíduos em duas classes, conforme segue:

Tabela 1: Dados para classificação da Rede Neural Artificial

INDIVÍDUO	JOVEM	IDOSO
MANOEL	X	
MARCO	X	
JOAO		X
JOSE		X

Fonte: O autor (2021)

O primeiro passo é codificar as informações em base binária, ou seja, 0 ou 1, a fim de que a nossa rede neural possa “entender” os valores a serem processados. A tabela abaixo representa os valores codificados.

Tabela 2: Codificação binária dos Dados para classificação da Rede Neural Artificial

	Código Binário
MANOEL	00
MARCO	01
JOAO	10

<b>JOSE</b>	11
<b>JOVEM</b>	0
<b>IDOSO</b>	1

Fonte: O autor (2021)

Conclui-se que apenas dois elementos na camada de entrada são necessários, pois o número de elementos necessários, em base binária, será dado pela fórmula:  $X = 2^n$ , em que  $n$  é o número de elementos na camada de entrada e  $X$  é o número de sinais de entrada. Ou seja, como temos quatro sinais de entrada, necessitaremos de dois neurônios nesta camada, pois ao aplicarmos a fórmula teremos  $4 = 2^2$ .

Para a camada de saída vale a mesma relação, ou seja, temos duas saídas previstas “Jovem” e “Idoso”, onde  $X$  é o número dos sinais de saída e neste caso igual a 2. Então a fórmula será  $X = 2^n$ , e  $X = “2”$ , e  $n = “1”$  que representa o número de elementos da camada de saída.

Com o objetivo de melhorar a capacidade de aproximação da rede neural será acrescentado um *bias* de entrada ao neurônio, pois, desta forma, haverá mais um parâmetro livre para ajustar. O sinal de entrada do *bias* “b” é constante e tem sempre o valor positivo 1, assim como a taxa de aprendizado será arbitrada para  $\eta = 1$ .

Como citado anteriormente, o cálculo do valor da função de ativação será efetuado pela equação:  $v = \sum_{i=1}^n w_i x_i$ , e a função de transferência escolhida terá a seguinte forma:

$$\varphi(v) = \begin{cases} 1, \forall v > 0 \\ 0, \forall v \leq 0 \end{cases}$$

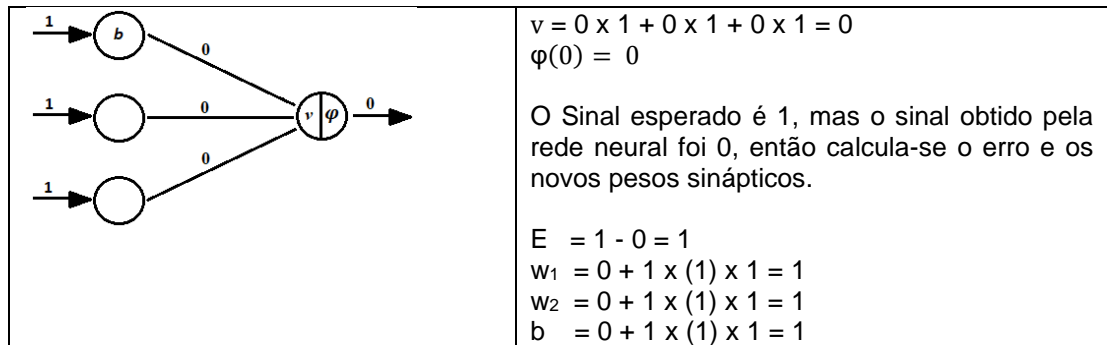
O cálculo dos novos pesos sinápticos e para o *bias*, a cada iteração ( $T$ ), segue a função  $w(i)_{T+1} = w(i)_T + \eta E_T x(i)$ , sendo  $i$  o neurônio respectivo.

Após estas considerações, iniciaremos o treinamento atribuindo o valor zero para todos os pesos sinápticos e para o *bias*, e exibiremos quadro a quadro todas as iterações durante esta fase.

Os quadros serão compostos do sinal de entrada, da imagem da rede neural, e dos cálculos da iteração. E para iniciar o treinamento será apresentado à rede neural o sinal JOSE (11).

Quadro 2: 1ª Iteração – Sinal de entrada JOSE (11)

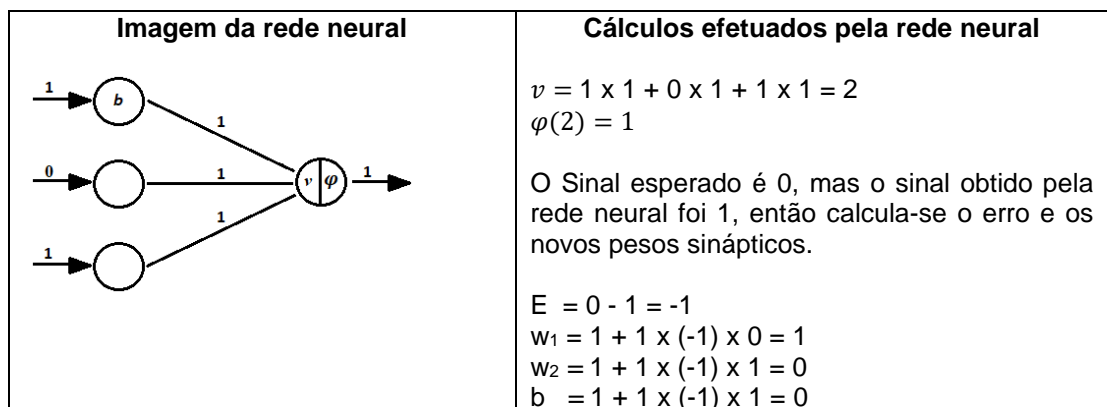
<b>Imagem da rede neural</b>	<b>Cálculos efetuados pela rede neural</b>
------------------------------	--



Fonte: O autor (2021)

Depois de efetuados os cálculos dos novos pesos sinápticos, em virtude do erro de classificação na 1ª iteração, apresentaremos para a rede neural o sinal MARCO (01), obtendo o quadro abaixo.

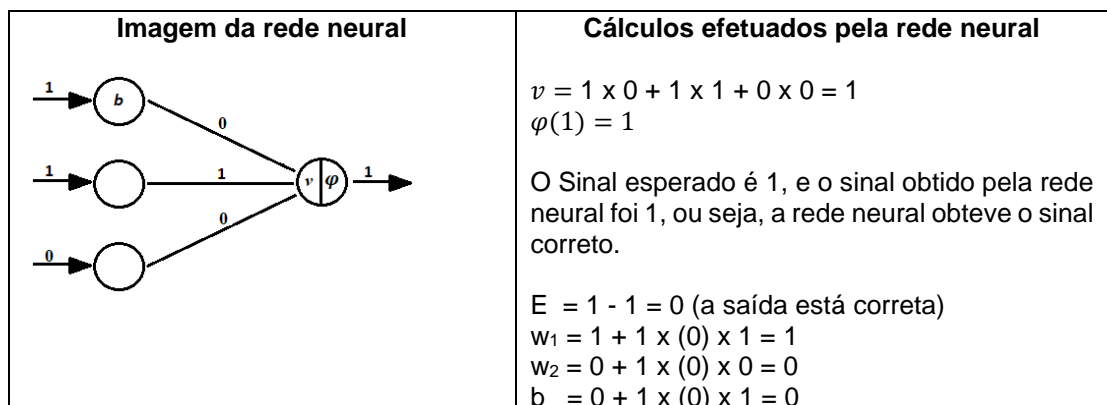
Quadro 3: 2ª iteração – Sinal de entrada MARCO (01)



Fonte: O autor (2021)

Prosseguindo com o treinamento, depois de efetuados os cálculos dos novos pesos sinápticos, em virtude do erro de classificação na 2ª iteração, apresentaremos para a rede neural o sinal JOAO (10), obtendo o quadro abaixo.

Quadro 4: 3ª iteração – Sinal de entrada JOAO (10)



Fonte: O autor (2021)

Nesta iteração podemos observar que a rede neural efetuou o cálculo corretamente, e por conseqüência os pesos sinápticos não sofrerão alterações nos seus valores. Prosseguindo com o treinamento, apresentamos o sinal MANOEL (00), conforme quadro abaixo.

Quadro 5: 4ª Iteração – Sinal de entrada MANOEL (00)

Imagem da rede neural	Cálculos efetuados pela rede neural
	$v = 1 \times 0 + 0 \times 1 + 0 \times 0 = 0$ $\varphi(0) = 0$ <p>O Sinal esperado é 0, e o sinal obtido pela rede neural foi 0, ou seja, a rede neural obteve o sinal correto</p> $E = 0 - 0 = 0 \text{ (a saída está correta)}$ $w_1 = 1 + 1 \times (0) \times 0 = 1$ $w_2 = 0 + 1 \times (0) \times 0 = 0$ $b = 0 + 1 \times (0) \times 1 = 0$

Fonte: O autor (2021)

Nesta iteração podemos observar que a rede neural, novamente, efetuou o cálculo corretamente, e por conseqüência os pesos sinápticos também não sofrerão alterações nos seus valores. Prosseguindo com o treinamento, rerepresentamos o sinal JOSE (11), conforme quadro abaixo.

Quadro 6: 4ª Iteração – Sinal de entrada JOSE (11)

Imagem da rede neural	Cálculos efetuados pela rede neural
	$v = 1 \times 0 + 1 \times 1 + 1 \times 0 = 1$ $\varphi(1) = 1$ <p>O Sinal esperado é 1, e o sinal obtido pela rede neural foi 1, ou seja a rede neural obteve, o sinal correto.</p> $E = 0 - 0 = 0 \text{ (a saída está correta);}$ $w_1 = 1 + 1 \times (0) \times 1 = 1$ $w_2 = 0 + 1 \times (0) \times 1 = 0$ $b = 0 + 1 \times (0) \times 1 = 0$

Fonte: O autor (2021)

Novamente observamos que a rede neural se comportou conforme esperado, efetuando a classificação correta, e ao rerepresentarmos para a rede neural todos os sinais de entrada obtemos erro médio igual a zero, significando que esta rede neural “aprendeu” a fazer a classificação desejada. O processo completo das iterações pode ser visto na tabela abaixo:

Tabela 3: Tabela das iterações do processamento da Rede Neural Artificial

iteração (T)	1	2	3	4	5	6	7
x1 Entradas	1	0	1	0	1	0	1
x2 Entradas	1	1	0	0	1	1	0
y1 esperado	1	0	1	0	1	0	1
y1 rede neural	0	1	1	0	1	0	1
Erro 1	1	-1	0	0	0	0	0
$v_1$	0	1	1	0	1	0	1
$w_{11}$	0	1	1	1	1	1	1
$w_{21}$	0	1	0	0	0	0	0
$w_{b1}$	0	1	0	0	0	0	0

Fonte: O autor (2021)

### 3.1.2 SEGUNDA APLICAÇÃO

Esta segunda aplicação apresenta uma rede neural Perceptron com duas entradas e duas saídas para mostrar como se comporta uma rede neural deste tipo um pouco mais complexa. Após seu treinamento, ela deverá ser capaz de classificar quatro indivíduos na forma a seguir:

Tabela 4: Dados para classificação da Rede Neural Artificial

	ESPAÑHOL	INGLES	JOVEM	IDOSO
JOSE		X	X	
JOAO	X		X	
HELENA	X			X
MARIA		X		X

Fonte: O autor (2021)

Seguindo os passos aplicados no exemplo anterior, codificam-se as informações em base binária (0 ou 1), as quais determinam-se como sendo:

Tabela 5: Codificação binária dos Dados para classificação da Rede Neural Artificial

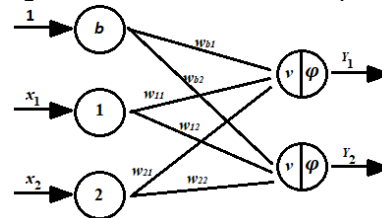
	Código Binário
JOSE	11
JOAO	10
HELENA	00
MARIA	01
ESPAÑHOL	0
INGLES	1
JOVEM	0
IDOSO	1

Fonte: O autor (2021)

Assim, serão utilizados apenas dois elementos na camada de entrada e dois neurônios na camada de saída. Assim como no exemplo anterior será incluído um *bias* de valor sempre positivo 1, cujos pesos são parâmetros livres para ajustamento, e a mesma taxa de aprendizado,  $\eta = 1$ .

O diagrama da rede neural é traçado abaixo, sendo aplicado inicialmente o valor zero para todos os pesos sinápticos e para o *bias*, exatamente como no exemplo anterior.

Figura 4: Rede Neural Perceptron



Fonte: LUDWIG e COSTA (2007, p. 37)

Para iniciar o treinamento desta rede neural, apresenta-se o sinal MARIA (01) à rede neural e esta retornará o sinal JOVEM ( $y_1 = 0$ ) e ESPANHOL ( $y_2 = 0$ ), respectivamente. As funções de ativação dos neurônios encontram-se descritas abaixo:

$$v_1 = 0 \times 1 + 0 \times 0 + 0 \times 1 = 0.$$

$$v_2 = 0 \times 1 + 0 \times 0 + 0 \times 1 = 0.$$

Dessa forma, a função de transferência utilizada (que é a mesma do exemplo anterior) retornará os seguintes valores de saída:  $\varphi_1(0) = 0$  e  $\varphi_2(0) = 0$ . E como as saídas não estão corretas, comparando-se com as saídas desejadas IDOSO = 1 e INGLES = 1, ter-se-ão os valores dos erros como:  $E_1 = 1 - 0 = 1$  e  $E_2 = 1 - 0 = 1$ , e os novos pesos serão:

- $w_{11} = 0 + 1 \times 1 \times 0 = 0$ ;
- $w_{12} = 0 + 1 \times 1 \times 0 = 0$ ;
- $w_{21} = 0 + 1 \times 1 \times 1 = 1$ ;
- $w_{22} = 0 + 1 \times 1 \times 1 = 1$ ;
- $w_{b1} = 0 + 1 \times 1 \times 1 = 1$ ;

- $w_{b2} = 0 + 1 \times 1 \times 1 = 1$ .

O cálculo dos novos pesos sinápticos a cada iteração (T) segue a função a  $w(i, j)_{T+1} = w(i, j)_T + \eta E(j)_T x(i)$ . O mapa completo de todas as iterações no treinamento está demonstrado abaixo.

Tabela 6: Tabela das iterações do processamento da Rede Neural Artificial

Iteração (T)	1	2	3	4	5	6	7	8	9	**	**	**	**
<b>x<sub>1</sub> Entrada</b>	0	0	1	1	0	0	1	1	0	0	0	1	1
<b>x<sub>2</sub> Entrada</b>	1	0	1	0	1	0	1	0	1	0	1	0	1
<b>y<sub>1</sub> Esperado</b>	1	1	0	0	1	1	0	0	1	1	1	0	0
<b>y<sub>2</sub> Esperado</b>	1	0	1	0	1	0	1	0	1	0	1	0	1
<b>y<sub>1</sub> Rede neural</b>	0	1	1	0	0	1	1	0	0	1	1	0	0
<b>y<sub>2</sub> Rede neural</b>	0	1	1	0	1	0	1	0	1	0	1	0	1
<b>Erro 1</b>	1	0	-1	0	1	0	-1	0	1	0	0	0	0
<b>Erro 2</b>	1	-1	0	0	0	0	0	0	0	0	0	0	0
<b>v<sub>1</sub></b>	0	1	1	0	0	1	1	0	0	1			
<b>v<sub>2</sub></b>	0	1	1	0	1	0	1	0	1	0			
<b>w<sub>11</sub></b>	0	0	0	-1	-1	-1	-1	-2	-2	-2			
<b>w<sub>12</sub></b>	0	0	0	0	0	0	0	0	0	0			
<b>w<sub>21</sub></b>	0	1	1	0	0	1	1	0	0	1			
<b>w<sub>22</sub></b>	0	1	1	1	1	1	1	1	1	1			
<b>w<sub>b1</sub></b>	0	1	1	0	0	1	1	0	0	1			
<b>w<sub>b2</sub></b>	0	1	0	0	0	0	0	0	0	0			

Fonte: O autor (2021)

Observemos que nas iterações 4, 6 e 8, os erros das duas saídas dos neurônios foram zero. Entretanto, na iteração posterior o erro novamente se apresenta no teste, e um novo cálculo dos pesos é requerido.

Ao atingirmos a décima iteração, as saídas dos neurônios tornam-se iguais às saídas esperadas, tornando todos os erros iguais a zero para qualquer um dos sinais de entrada, ou seja, a rede neural está treinada.

## 4 CONSIDERAÇÕES FINAIS

É inegável a evolução dos algoritmos de Aprendizado de Máquina ao longo das últimas décadas, assim como o fato destes algoritmos terem se tornado uma das grandes promessas de evolução nas mais diversas tarefas do nosso cotidiano.



Dentre tantos algoritmos as redes neurais ganharam grande destaque, muito em parte devido a sua capacidade de aprendizado, por intermédio de treinamento adequado, simulando uma das capacidades humanas.

Neste trabalho o intuito foi o de exemplificar a matemática que está por trás das redes neurais, e para isso foram utilizados dois exemplos bem simples, e uma rede neural de igual simplicidade, mas que serviram ao propósito.

Cabe ressaltar que em todos os algoritmos de Aprendizado de Máquina a matemática está presente de forma significativa, e decisiva, sempre servindo de base para os desenvolvimentos computacionais que virão a seguir.

Contudo, devemos ter a plena consciência de que a utilização destes algoritmos, que hoje podem ser feitos através de inúmeras bibliotecas com os códigos prontos para utilização (*frameworks*), não determinam o processo como um todo, sendo essa apenas uma etapa do processo de Inteligência Artificial.

Outra abordagem para um trabalho futuro, seria ilustrar as etapas de preparação dos dados para implementação dos algoritmos, pois nestas etapas há que se observar as correlações entre as variáveis, tratamento de valores nulos, tratamento dos *outliers*, e nessas tarefas a presença da matemática é decisiva.

## REFERÊNCIAS

FONSECA, J. J. S. **Metodologia da pesquisa científica**. Fortaleza: UECE, 2002.

HAYKIN, Simon. **Redes Neurais: princípios e prática**. trad. Paulo Martins Engel. 2 ed.;Porto Alegre. Bookman. 2001.

LAUDON, Kenneth C.; LAUDON, Jane P.. **Management Information Systems**. Prentice Hall. 1996.

LUDWIG JR, O; COSTA, Eduard Montgomery M. **Redes Neurais - Fundamentos e Aplicações com Programas em C**. Rio de Janeiro. Editora Ciência Moderna Ltda. 2007.

MITCHELL, Tom M.. **Machine Learning**. McGraw-Hill, 1. Ed, 1997.

ROWE, N. C.. **Artificial intelligence through Prolog**. Prentice-Hall, Englewood Cliffs, New Jersey. 1988.

RUSSEL, Stuart; NORVIG, Peter. **Inteligência Artificial**. 3 ed. Rio de Janeiro: Elsevier, 2013.

SAXENA, Shubh. Artificial Neuron Networks(Basics) Introduction to Neural Networks. Disponível em <https://becominghuman.ai/artificial-neuron-networks-basics-introduction-to-neural-networks-3082f1dcca8c>. Acesso em 11 de janeiro de 2021.