



**UNIVERSIDADE ESTADUAL DA PARAÍBA  
CAMPUS I  
CENTRO CIÊNCIAS E TECNOLOGIA  
DEPARTAMENTO DE MATEMÁTICA  
CURSO DE LICENCIATURA EM MATEMÁTICA**

**JOSÉ MAX DE SOUZA ALVES**

**REDES NEURAS ARTIFICIAIS: UMA INTRODUÇÃO À MATEMÁTICA  
APLICADA**

**CAMPINA GRANDE  
2023**

JOSÉ MAX DE SOUZA ALVES

**REDES NEURAS ARTIFICIAIS: UMA INTRODUÇÃO À MATEMÁTICA  
APLICADA**

Trabalho de Conclusão de Curso (Artigo) apresentado ao Curso de Licenciatura em Matemática –CCT – UEPB, como requisito parcial para obtenção do título de Licenciado em Matemática.

**Área de concentração:** Matemática Aplicada

**Orientador:** Profa. Dra. Emanuela Régia de Sousa Coelho

**CAMPINA GRANDE  
2023**

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

A474r Alves, Jose Max de Souza.  
Redes neurais artificiais [manuscrito] : uma introdução à matemática aplicada / Jose Max de Souza Alves. - 2023.  
27 p. : il. colorido.

Digitado.  
Trabalho de Conclusão de Curso (Graduação em Matemática) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2023.  
"Orientação : Prof. Dr. Emanuela Régia de Sousa Coelho, Coordenação do Curso de Matemática - CCT. "

1. Redes neurais. 2. Perceptron. 3. Neurônios. 4. Treinamento de máquina. 5. Inteligência artificial . I. Título  
21. ed. CDD 006.32

JOSÉ MAX DE SOUZA ALVES

Redes Neurais Artificiais: uma introdução à Matemática Aplicada

Trabalho de Conclusão de Curso (Artigo)  
apresentado ao Curso de Licenciatura em  
Matemática –CCT – UEPB, como requi-  
sito parcial para obtenção do título de Li-  
cenciado em Matemática.

Área de concentração: Matemática Apli-  
cada

Aprovado em: 07 /09/2023.

**BANCA EXAMINADORA**

---

Profa. Dra. Emanuela Régia de Sousa Coelho  
Universidade Estadual da Paraíba (UEPB)

---

Prof. Me. Jair Dias de Abreu  
Universidade Estadual da Paraíba (UEPB)

---

Prof. Me. Renan Pires de Araújo  
Universidade Estadual da Paraíba (UEPB)

*Dedico este trabalho a Deus; sem ele eu não teria capacidade para desenvolver este trabalho. Dedico aos meus pais, pois é graças ao seus esforços que hoje posso concluir o meu curso.*

“Se a educação sozinha não transforma a sociedade, sem ela tampouco a sociedade muda.”

Paulo Freire

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>6</b>
<b>2</b>	<b>REDES NEURAIS: ONDE COMEÇA TUDO</b>	<b>8</b>
<b>2.1</b>	Neurônios	8
<b>2.2</b>	Sistema Nervoso	10
<b>2.2.1</b>	Sistema Nervoso Central	11
<b>2.2.2</b>	Sistema Nervoso Periférico	12
<b>3</b>	<b>REDES NEURAIS ARTIFICIAIS</b>	<b>13</b>
<b>3.1</b>	Principais Redes Neurais	13
<b>3.1.1</b>	Arquitetura Feedforward mais simples	14
<b>4</b>	<b>TREINAMENTO DA REDE NEURAL</b>	<b>15</b>
<b>4.1</b>	Treinamento sob supervisão	16
<b>4.2</b>	Treinamento sem supervisão	17
<b>4.3</b>	Perceptron da Camada Única	17
<b>5</b>	<b>ONDE A MATEMÁTICA ENTRA: O PERCEPTRON</b>	<b>18</b>
<b>5.1</b>	Convergência do Perceptron	19
<b>6</b>	<b>APLICAÇÃO</b>	<b>21</b>
<b>7</b>	<b>CONCLUSÃO</b>	<b>24</b>
	<b>REFERÊNCIAS</b>	<b>25</b>

## REDES NEURAIS ARTIFICIAIS: UMA INTRODUÇÃO À MATEMÁTICA APLICADA

Jose Max de Souza Alves\*

### RESUMO

Cada dia mais cientistas e estudiosos buscam meios e mecanismos para tentar compreender melhor a mente humana a fim de entender melhor os pensamentos, as tomadas de decisões, as próprias ações e pensamentos frente a uma sociedade inconstante. Todo esse compilado de informações está vinculado previamente a uma unidade morfológica denominada neurônio que nos dá base para nosso estudo. O seguinte trabalho descreve o funcionamento da rede neural Perceptron, envolvendo os processos de aprendizado da máquina, das saídas dos valores adotados para seus pesos e bias, bem como sua aplicação em processos simples de separação e classificação, tudo isto possibilitado pelo uso da matemática aplicada.

**Palavras-chave:** redes neurais; perceptron; neurônios; treinamento da máquina.

### ABSTRACT

Scientists and scholars increasingly seek means and mechanisms to try to better understand the human mind in order to better understand thoughts, decision-making, actions and thoughts in the face of an unstable society. All this compilation of information is previously linked to a morphological unit called neuron that gives us the basis for our study. The aim of this work is to describe the operation of the Perceptron neural network, involving machine learning processes, the output values adopted for their weights and bias, besides their application in simple processes of separation and classification, all of which made possible by the use of applied mathematics present in this research.

**Keywords:** neural networks; perceptron; neurons; machine training.

## 1 INTRODUÇÃO

Uma das estruturas biológicas mais fascinantes e intrigantes existentes em todos os seres humanos, palco de diversos estudos nas mais diversas áreas e que por muitas vezes é incompreendido, pesa por volta de 1,36 Kg e com 100 bilhões de neurônios, possuindo diversos processos de altíssima complexabilidade: estamos falando do cérebro.

Neste estudo buscamos compreender duas questões que norteiam este trabalho: sabemos que todo o corpo é formado por células das mais variadas magnitudes e funções dentro do nosso sistema, então qual seria a célula responsável por gerar os pensamentos humanos; e assim sendo este trabalho na área da matemática aplicada, qual seria a rede neural que copia esse funcionamento desta célula.

E nesse momento nos pegamos a pensar o porque do cérebro intrigar tanto, pois queremos ter o entendimento de como que várias células interconectadas podem formar cada indivíduo com seus traços de caráter, pensamentos lógicos racionais e personalidade única de cada pessoa.

---

\*Aluno de graduação do Curso de Licenciatura em Matemática do Departamento de Matemática, da Universidade Estadual da Paraíba. E-mail:jose.max@aluno.uepb.edu.br. Este artigo de conclusão de curso foi escrito sob orientação da Profa. Dra. Emanuela Régia de Sousa Coelho.



Pensamentos, ações, movimentos, observações, tudo parte dos comandos recebidos, absorvidos, entendidos e replicados pelo e no cérebro através dos mais de 100 bilhões de neurônios que são responsáveis pelo processamento de informações e que formam essa verdadeira máquina de natureza humana. Este órgão tem várias funcionalidades de conexões e controle corporal. Podemos entender também o cérebro humano como um computador altamente complexo, não linear e paralelo (HAYKIN, 2001).

Por exemplo, peguemos o sistema respiratório responsável pela percepção dos aromas e cheiros, bem como manter a oxigenação necessária à nossa sobrevivência, podemos entender neste estudo que muitos dos comandos viram de sistemas, sistemas estes denominados de sistema nervoso autônomo ou neurovegetativo, através de um centro nervoso localizado na região do Bulbo.

Algo semelhante acontece nos animais, por exemplo, na respiração dos cães.

O centro respiratório, composto de diversos grupos de neurônios localizados bilateralmente no bulbo e na ponte, desempenha o papel de controle da ventilação. Esses grupos enviam impulsos para os músculos respiratórios (diafragma e intercostais), estimulando contração dos mesmos. (Dukes, 1996, p. 54)

Mas como se dá todos esses processos? Na formação do cérebro temos o chamado "processo de aprendizado" o qual é efetivo para a formação das ações através de experiências contundentes, tudo proporcionado pela neuroplasticidade do cérebro em seu estágio inicial de formação.

Tudo isso a uma velocidade de processamento muita mais rápido que qualquer computador com a mais alta tecnologia já desenvolvida. Por isso, o homem sentiu-se na necessidade de buscar meios para entender o funcionamento da mente humana através de máquina, tentar ensiná-los através dos processos de aprendizagem, que consistem em complexos mecanismos de dados os quais são convertidos e processados de acordo com a função da máquina.

Todo esse processo podemos chamar de algoritmo da aprendizagem, cuja função é modificar os pesos sinápticos da rede de uma forma ordenada para alcançar um objetivo de projeto desejado (HAYKIN, 2001) e, por fim, termos o comportamento desejado através das várias sinapses que existem nessa adorável máquina chamada cérebro.

Para a neurociência, a aprendizagem, que faz parte dos processos mentais superiores, é toda forma que o indivíduo tem de adquirir conhecimento de forma física ou cognitiva que acontece com ele através do decorrer de sua vida. Para Gazzaniga, Ivry e Mangun (2006, p.28), temos que "aprendizagem é o processo de aquisição de informação, enquanto memória refere-se à persistência do aprendizado em um estado que pode ser evidenciado posteriormente".

Assim, temos que a neurociência traz em si uma responsabilidade fundamental que é a de compreender as bases biológicas da consciência e dos processos mentais superiores por meio dos quais percebemos, agimos, aprendemos e memorizamos eventos.

O meio social é de extrema importância para que a interação no sistema nervoso funcione, pois é através dos estímulos que boa parte das experiências são vivenciadas pelas pessoas.

Fato importante quando se fala de neurociência e os processos de aprendizado é que estes não cessam com o passar do tempo, simplesmente passam a ter focos em outros eventos, abrangendo uma maior gama de afirmação e conseqüentemente trazendo mais conhecimentos para o indivíduo.

O estudo da neurociência deriva dos estudos de diferentes áreas da ciência: medicina, biologia, psicologia, física, química e matemática. Os pensamentos primordiais da neurociência surgiram com Hipócrates (460 - 379 a.C.), o pai da medicina ocidental, que considerava o encefalo como sendo a encéfalo seria a central da inteligência, ideia esta que não foi bem aceita por todos.

Já Aristóteles (384 - 320 a. C.) tinha para si que o cérebro era o centro de tudo e o encéfalo era o controlador da pressão sanguínea. Entretanto, Galeno (130 - 200 d. C.) concordava com os pensamentos de Hipócrates. Suas ideias e descobertas perduraram por quase 1500 anos.

Andreas Vesalius (1512 - 1564) em seus estudos aderiu mais detalhes à estrutura encefálica elaborada por Galeno, pois Vesalius ia em contraste com as ideias de que o coração era o centro das emoções e da mente, sendo definidas estas ao cérebro, isto pelo fato dos nervos serem originários do mesmo, e não do coração.

Já René Descartes (1596 - 1650) acreditava que o encéfalo era responsável pelo comportamento humano enquanto que a consciência estaria sendo produzida fora da "mente".

Com todas essas descobertas, a neurociência foi evoluindo até o ponto de determinar todas as partes constituintes do cérebro, o qual este trabalho esta debruçado inicialmente.

Por fim, vamos apresentar neste trabalho todas as estruturas responsáveis pelo funcionamento do cérebro humano. Além disto, iremos apresentar o conceito de processo de aprendizado das Redes Neurais Artificiais (RNA's) e sua aplicação final na rede neural Perceptron de camada única, tudo isto fazendo uso da matemática aplicada.

## 2 REDES NEURAIIS: ONDE COMEÇA TUDO

Nessa seção, apresentamos um pouco da base do desenvolvimento de Redes Neurais Artificiais: o funcionamento do Cérebro Humano.

### 2.1 Neurônios

Podemos dizer que o título desta seção é o que dá sentido aos estudos deste trabalho. Segundo Catarina Moreira (2013), o neurônio:

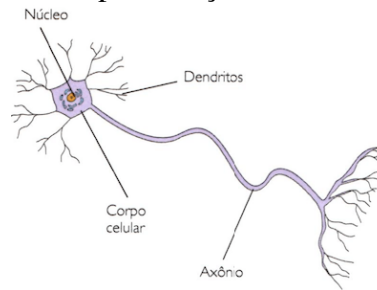
É uma célula nervosa, estrutura básica do sistema nervoso, comum à maioria dos vertebrados. Os neurônios são células altamente estimuláveis, que processam e transmitem informação através de sinais eletro-químicos. Uma das suas características é a capacidade das suas membranas plasmáticas gerarem impulsos nervosos. (Moreira, 2013, p. 24)

Também chamada de célula nervosa, é a estrutura base para o bom funcionamento do sistema nervoso, sendo responsável por ser a receptora dos estímulos dos altos níveis de interações das células do nosso corpo, além de ser responsável também por processar todas as informações e enviar novos estímulos ao nosso organismo.

A maioria dos neurônios geralmente tem sua estrutura formada pelo corpo celular e dois tipos de extensões e dois tipos de extensões citoplasmáticas, dendritos e axônios, em que

- Corpo celular: contém o núcleo e a maior parte das organelas celulares. É aqui que a composição de proteína acontece.
- Dendritos: São pequenas extensões, geralmente ramificação, recebem e conduzem estímulos de outros neurônios ou células do sentido.
- Axônio: é a extensão, geralmente a mais longa, que transmite os impulsos nervosos do corpo celular. Os axônios variam em comprimento entre muitos tipos diferentes de neurônios.
- Terminais Axônicos: contêm sinapses, estruturas especializadas que liberam substâncias químicas, chamadas neurotransmissores, além de fazerem conexões com dendritos ou com o corpo celular de outros neurônios.

Figura 1: Representação de um Neurônio



Fonte: AMABIS; MARTHO (2006)

Quando o terminal axônico de um neurônio se conecta com um dendrito ou corpo celular de outro neurônio, a membrana muda e uma sinapse é formada, o que permite que os impulsos nervosos viajem de um neurônio para o próximo. Quando o impulso nervoso chega ao fim, os neurotransmissores axônicos que formam as sinapses são liberados da membrana pré-sináptica, atravessam a fenda sináptica e se ligam aos receptores da membrana pós-sináptica do próximo neurônio.

No entanto, os neurônios não são as únicas células do cérebro. No sistema nervoso, a função das células gliais e os neurônios são de se conectar para ajudar no desenvolvimento embrionário.

Há alguns tipos de células gliais: células de Schwann no sistema nervoso periférico e oligodendrócitos no sistema nervoso central. Muitas células gliais fornecem nutrientes para os neurônios, enquanto outros consomem partículas estranhas e detritos celulares.

Outro função é manter os níveis de íons ao redor dos neurônios. Embora careçam de axônios para a transmissão de impulsos nervosos, as células gliais comunicam-se eletricamente uns com os outros através de "gap", "junções", que permitem o fluxo de íons entre as células.

Como todas as células, o citoplasma dos neurônios com carga negativa. A voltagem dentro de um neurônio é tipicamente 60-70 milivolts (mV) mais negativo do que extracelular. Esta diferença de carga entre os ambientes extracelular e intracelular cria uma diferença de potencial entre as duas superfícies da membrana - o potencial de membrana onde quando a célula não está transmitindo impulsos nervosos é da ordem de -70 mV potencial de repouso.

Os neurônios são sensíveis aos estímulos causando assim as mudanças químicas ou físicas potenciais das membrana em repouso ocasionando maior mudança dos impulsos nervosos (ou potenciais de ação), que é a rápida mudança no potencial, onde brevemente entre 1 ou 2 milissegundos dentro as células, tornam-se mais agressivas ou reativas ao exterior.

A membrana plasmática dos neurônios é composta por uma bicamada fosfolipídica impermeável à água e a íons, como outras células, mas eles têm proteínas e atua como um canal iônico ou bomba.

O potencial de repouso é principalmente devido à diferença entre a concentração de íons sódio  $Na^+$  e potássio  $K^+$  dentro e fora da célula. Essa diferença é causada através do canal de sódio juntamente com a atuação da bomba e potássio, que bombeia sódio para o meio externo e potássio para o meio interno, consome ATP e neutraliza a difusão passiva desses íons.

A bomba de sódio e potássio fornece 3  $Na^+$  por quantidade de íons  $K^+$  deixando a célula por 2  $K^+$  (via transporte passivo) maior que íons  $Na^+$  entrando na célula, causando carga positiva na bateria em relação ao exterior.

Canais presentes na membrana celular permitem que  $K^+$  e  $Na^+$  passem passivamente: quando o neurônio está em repouso, o canal está fechado; quando a célula é estimulada, o canal é aberto, permitindo que  $Na^+$  entre rapidamente, causando mudanças potenciais na membrana de -70 mV a +35 mV, chamamos esta diferença de potencial de despolarização. Assim a rápida

variação de potencial que ocorre durante a despolarização é chamada de potencial de ação e é da ordem de 105 mV.

Quando o valor máximo do potencial de ação é alcançado durante a despolarização, a permeabilidade ao  $K^+$  da membrana aumenta e a permeabilidade do canal de  $Na^+$  volta ao normal.

A transmissão de impulsos nervosos é um exemplo de resposta "tudo ou nada", ou seja, o estímulo deve ter uma certa intensidade e gerar potencial de ação. A estimulação mínima necessária para desencadear o potencial de ação é o estímulo limiar.

Uma vez que este limiar é atingido, o aumento da força não cria potencial e nem ação mais forte, mas mais pulsos por segundo.

O potencial de ação gerado na membrana estimulada se propaga para regiões adjacentes, causando sua despolarização. Essa despolarização e repolarização contínuas leva a membrana de um neurônio gera um impulso nervoso, cuja propagação acaba se tornando uma direção dos dendritos para os axônios.

Diferentes velocidades de transmissão de impulsos nervosos existem entre neurônios de espécies diferentes.. Por exemplo, em geral, a velocidade dos anêmonas do mar é 0,1 m/s, enquanto nos neurônios motores de alguns mamíferos a velocidade é de cerca de 120m/s. A diferença nessas velocidades de transferência está relacionada com a estrutura do axônio:

- Diâmetro: pequeno diâmetro apresenta grande resistência e faz com que os impulsos viajem mais lentamente;
- Mielina: em vertebrados, embora os axônios sejam menores em diâmetro do que nos invertebrados e viajem mais rápido, o impulso é garantido pela presença da bainha mielina, formada por células de Schwann que envolvem os axônios.

Nos axônios mielinizados, apenas o potencial de ação despolariza a membrana na área nodular Ranvier, pois a bainha atua como isolante, impedindo a despolarização no restante da área.

Quando pulsado de um nó para outro, os impulsos nervosos são transmitidos de uma célula para outra que é feito através de sinapses.

## 2.2 Sistema Nervoso

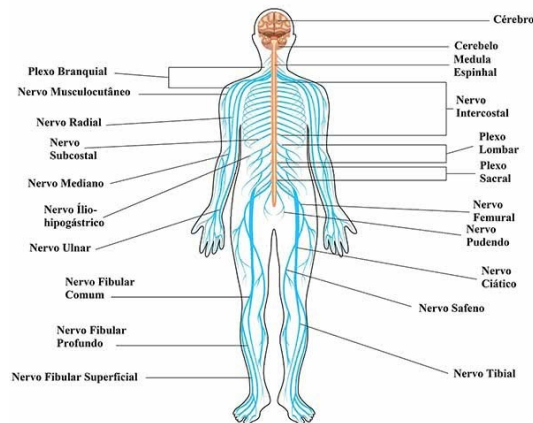
O sistema nervoso representa uma rede de comunicações do organismo o qual é responsável pela execução de diversas atividades e seu entendimento através do raciocínio lógico além das percepções do mundo que cerca o indivíduo.

Este é constituído por todo um conjunto de órgãos do corpo humano que tem a função de receber as mensagens ou os estímulos do ambiente, "interpretá-los" e "arquivá-los" de acordo com cada função ao área que o sistema estiver trabalhando. Conseqüentemente, ele elabora respostas, as quais podem ser dadas na forma de movimentos, sensações ou constatações.

O sistema nervoso tem três funções básicas:

- Sensível: é quando vários estímulos e informações são captados por receptores sensitivos, também chamados de neurônios receptores, localizados em todo o corpo. Por exemplo, uma lesão na pele ou um aumento na temperatura externa podem fazer com que os neurônios sensoriais disparem e levem esses estímulos até o cérebro, transformando em mensagem para que possamos nos afastar de algo que venha ser prejudicial.
- Integradores: acontece quando as células nervosas, chamadas de interneurônios ou neurônios conectivos, analisam, processam e armazenam estímulos e informações captadas por

Figura 2: Sistema Nervoso Central



Fonte: BURT (1995)

receptores sensoriais, o chamado "arquivamento" de informações, por exemplo, em momentos de estudo.

- Movimento: a etapa final é realizada pelos neurônios motores, também conhecidos como neurônios eferentes ou efetores, que fazem contato com órgãos efetores, recebem informações do cérebro e realizam ações de acordo com a situação.

O Sistema Nervoso está dividido em duas partes fundamentais: sistema nervoso central e sistema nervoso periférico.

### 2.2.1 Sistema Nervoso Central

O Sistema Nervoso Central, também chamado de sistema neural, é constituído pelo encéfalo e pela medula espinhal, ambos envolvidos por três membranas denominadas meninges. O encéfalo está localizado inteiramente no crânio enquanto que a medula espinhal localiza-se na coluna cervical.

Ele fica responsável pelo controle das ações voluntárias como, por exemplo, falar, andar como pelas ações involuntárias como a digestão dos alimentos, a respiração, os batimentos cardíacos.

#### ENCÉFALO

O encéfalo, que pesa aproximadamente 1,5 Kg, está localizado na caixa craniana e apresenta três órgãos principais: o cérebro, o cerebelo e o tronco encefálico.

#### CÉREBRO

Podemos considerar dentro de todo o processo de pesquisa das redes neurais artificiais o órgão mais importante do sistema nervoso. Considerado que seja o órgão mais volumoso, pois ocupa a maior parte do encéfalo, o cérebro está dividido em duas partes simétricas: o hemisfério direito e o hemisfério esquerdo. A camada mais externa do cérebro é cheia de dobraduras, chama-se córtex cerebral, o que possui as mais diversas funções sendo responsável pelo pensamento, visão, audição, tato, paladar, fala, escrita. Ademais, é sede dos atos conscientes e inconscientes, da memória, do raciocínio, da inteligência e da imaginação, e controla ainda os movimentos voluntários do corpo.

#### CEREBELO

Ele se encontra na parte posterior e abaixo do cérebro, possui grande parte dos neurônios e é responsável pela coordenação dos movimentos que precisamos para manter o corpo em movimento, além estabelecer o equilíbrio corporal. Regula o tônus muscular, ou seja, regula o grau de contração dos músculos em repouso.

### **TRONCO ENCEFÁLICO**

Este é responsável pelos impulsos nervosos do cérebro para a medula espinhal e vice-versa. É composto pelo mesencéfalo, pela ponte e pelo bulbo sendo sua forma semelhar a de uma calda. Ele possui uma importância tremenda pois é através dele que se produz os estímulos nervosos que servem para controlar as atividades vitais como os movimentos respiratórios, os batimentos cardíacos e os reflexos, como a tosse, o espirro e a deglutição.

### **MEDULA ESPINHAL**

Seguindo a ordem anatômica, a medula espinhal começa após o termino do tronco encefálico que é uma verdadeira fiação de tecidos nervosos situados por dentro da coluna vertebral. Sua finalidade é conduzir os impulsos nervosos de todo o corpo, desde os estímulos recebidos na pele, por exemplo, sobre situações que estão acontecendo fora do nosso corpo, até as informações involuntárias dos nossos órgãos, passando a coletar todas essas informações e levando até o cérebro para que este coordene o que se deve fazer com esses estímulos.

#### **2.2.2 Sistema Nervoso Periférico**

É importante trazermos o sistema nervoso periférico pois é por onde todas as informações vinda dos órgãos passam e vão ao Sistema Nervoso Central (SNC) e por fim ele conecta o sistema nervoso central ao resto do corpo. Nos estudos das Redes Neurais Artificiais (RNA) se assemelha ao sistema de aprendizado, onde ele recebe as informações, tendo de transformar em novos estímulos e alimentar o sistema. Importante destacar que existem dois tipos de nervos: os cranianos e os raquidianos. O sistema nervoso periférico é formado por nervos que se originam no encéfalo e na medula espinhal.

### **NERVOS CRANIANOS**

Distribuem-se em 12 pares que saem do encéfalo, e sua função é transmitir mensagens sensoriais ou motoras, especialmente para as áreas da cabeça e do pescoço.

**NERVOS RAQUIDIANOS** Podemos dizer que os nervos raquidianos ou nervos espinal são um emaranhado de 31 nervos que possuem diversas funções. Estes nervos são formado por neurônios que captam os estímulos que recebem do ambiente tem a função chamada mista porque podem desempenhar a função motora já que são um ramificado que saem da medula espinal como desempenham a função sensitiva captando os estímulos dos órgãos.

### **SISTEMA NERVOSO SOMÁTICO**

Regula as ações voluntárias, ou seja, que estão sob o controle da nossa vontade bem como regula a musculatura esquelética de todo o corpo.

### **SISTEMA NERVOSO AUTÔNOMO**

Este sistema é responsável pelo controle das funções viscerais do organismo, ou seja, controla a parte do nosso sistema motor e são responsáveis pelos músculos lisos, o músculo cardíaco, as glândulas e parte do tecido adiposo. Com isso desempenham uma importante parte do sistema do ambiente interno do nosso organismo além de captar estímulos necessários as respostas externas do nosso organismo.

### 3 REDES NEURAIIS ARTIFICIAIS

Discutiremos agora as várias topologias e projetos arquitetônicos de redes neurais artificiais que já existem e alguns dos seus métodos de treinamento. Para que tenhamos os estudos das redes neurais, é preciso entender como realmente funciona o cérebro humano, como foi discutido nos capítulos anteriores, pois as redes neurais “copiam” os mecanismos de funcionamento da mente humana. Através da arquitetura de uma rede neural, estamos examinando como os neurônios se acomodam enquanto mantêm as conexões sinápticas discutidas anteriormente. Por outro lado, a topologia da rede refere-se aos vários arranjos estruturais com números variados de neurônios nos nós de entrada, meio e saída da rede.

O processo de treinamento, também conhecido como processo de adaptação iterativa, é como o aprendizado de uma rede neural artificial é colocado em prática. O aprendizado finaliza apenas quando uma rede neural descobre uma solução amplamente aplicável a um problema específico. Concluindo, treinar uma rede implica ajustar sua matriz de pesos sinápticos para que o peso coincida com a resposta de saída da rede.

A aplicação de um conjunto de etapas ordenadas com o objetivo de ajustar o peso e as condições de contorno de uma determinada forma arquitetônica equivale ao exercício dessa arquitetura. Com isso, esse procedimento de ajuste, também conhecido como algoritmo de aprendizado, visa ajustar a rede para que as respostas fiquem próximas aos valores de referência assumidos.

Como resultado, alguns tipos diferentes de arquiteturas de redes neurais artificiais e seus nomes literários correspondentes serão discutidos nas seções a seguir.

#### 3.1 Principais Redes Neurais

A organização de redes neuronais artificiais é um fator crucial que precisa ser levado em consideração. Existe uma forte restrição no número de camadas que uma rede neural pode ter; como resultado, uma rede neural pode ter uma ou várias câmeras. A rede neural que possui uma única camada é aquela que tem um nó entre uma entrada e uma saída. Por outro lado, as redes neurais com várias camadas têm uma as camadas de entrada e saída do sistema.

Nas redes artificiais multicamadas, existem camadas ocultas, também conhecidas como intermediárias. Este número não especificado de camadas está situado entre a camada de entrada e a camada de saída da rede neural (CARVALHO e LUDEMIR, 1998). Segundo Silva, Spatti e Flauzino (2010), uma rede neural artificial pode ser dividida em três camadas:

1) Camada de entrada: Esta camada é responsável por receber dados, sinais e características externamente admitidas; muitas vezes essas entradas são normalizadas em relação às faixas de variações dinâmicas produzidas pelas funções de acionamento. Essa normalização resulta em maior precisão numérica nas operações matemáticas da rede.

2) Camadas ocultas ou intermediárias: são aquelas compostas por neurônios que têm a função de extrair as características relacionadas ao processamento interno da rede que é realizado nesta camada

3) Camada de saída: Esta camada também é formada por neurônios e se encarrega de produzir e apresentar os resultados finais da rede, que são os resultados das operações de processamento realizadas pelos neurônios nas etapas anteriores.

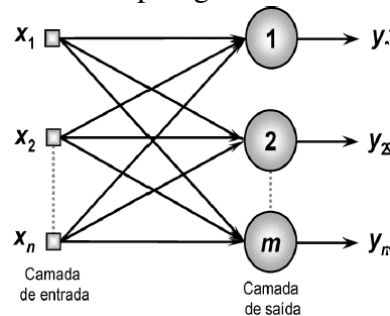
Um bom exemplo desse tipo de arquitetura são as redes *feedforward*, onde a mensagem é constantemente propagada da entrada à saída. Como um todo, as redes *feedforward* são organizadas em *clusters* e atualmente são muito utilizadas. Elas podem ser divididas em quatro categorias: redes de camada simples, redes de camada múltipla, redes recorrentes e redes reticuladas que nada mais é do que a rede que considera a posição espacial do neurônio para o

ajuste dos pesos e dos bias.

### 3.1.1 Arquitetura Feedforward mais simples

O fluxo de informação é sempre na direção as saídas que é onde iremos ter os resultados esperados ou não, e as seguintes redes neurais são usados para resolver problemas de classificação e filtragem: Perceptron e Adaline, cujos algoritmos de aprendizado são baseados nos regimes Hebb e Delta, respectivamente. A quantidade de entradas em redes relacionadas à arquitetura será, portanto, sempre igual ao número de neurônios.

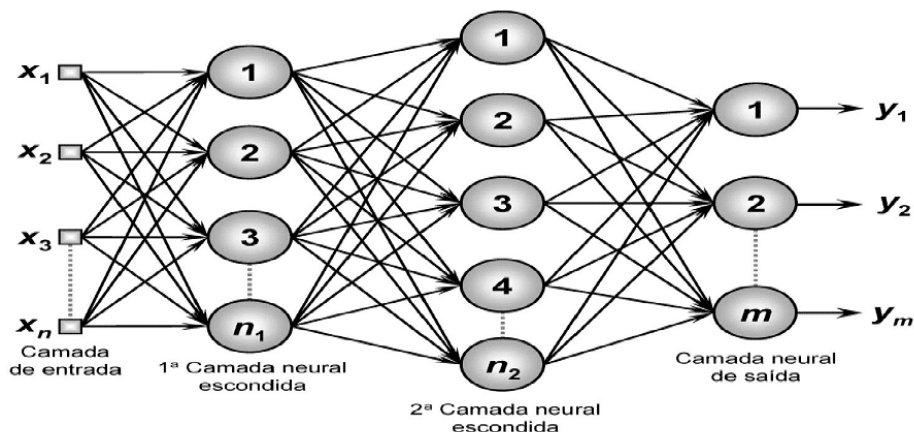
Figura 3: Arquiteturas e Topologias de Redes Neurais Artificiais



Fonte: Página Embarcados <sup>1</sup>

Uma rede *feedforward* de redes neurais ocultas é composta por uma ou mais dessas camadas. Eles são usados em problemas envolvendo solução aproximada de funções, classificação de parâmetros, identificação de sistemas, otimização, robótica e controle de processos. Essa rede é composta por uma camada de entrada composta por  $n$  sinais, duas camadas neurais ocultas compostas por  $n_1$  e  $n_2$  neurônios e uma câmara neural composta por  $m$  neurônios.

Figura 4: Arquiteturas e Topologias de Redes Neurais Artificiais



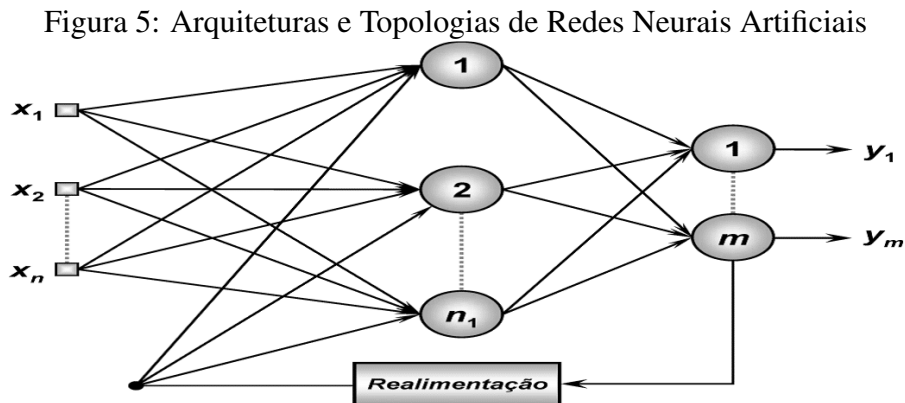
Fonte: Página Embarcados <sup>2</sup>

<sup>1</sup>Disponível em: [www.embarcados.com.br/redes-neurais-artificiais/](http://www.embarcados.com.br/redes-neurais-artificiais/). Acesso em 21 Novembro 2022

<sup>2</sup>Disponível em: [www.embarcados.com.br/redes-neurais-artificiais/](http://www.embarcados.com.br/redes-neurais-artificiais/). Acesso em 24 Novembro 2022



Segundo Silva, Spatti e Flauzino (2010), em relação a redes neurais com arquitetura *feed-forward* de camadas simples, é possível abstrair que o número de neurônios que compõem a primeira camada oculta é tipicamente diferente do número de sinais que compõem a camada de entrada da rede”. Na verdade, o número de camadas ocultas e os neurônios que cada uma contém depende do tipo e nível de complexidade do problema a ser mapeado na rede, bem como da quantidade e calibre dos dados disponíveis para resolver o problema.



Fonte: Página Embarcados <sup>3</sup>

As redes neurais artificiais mais generalizadas são aquelas que possuem varias camadas de abastecimento nas entradas, sendo suas camadas determinadas tanto pelas entradas atuais quanto pelas entradas anteriores. Assim, a estrutura desta rede não é necessariamente organizada em camadas, e se fosse, poderia haver conexões entre neurônios de uma mesma camada ou mesmo entre camadas não consecutivas. Como incluem recursos de realimentação, essas redes reagem às entradas dinamicamente, o que significa que, após a aplicação de uma nova entrada, é feito um cálculo de saída e a rede é realimentada para alterar a entrada. Por esse motivo, as redes são chamadas de "com memória".

Silva, Spatti e Flauzino (2010) enfatizam a ideia da arquitetura de redes neurais artificiais que se assemelham a redes nas quais as mortes de neurônios servem como sinais de entrada para outros neurônios. Tais redes são qualificadas para processamento dinâmico de informações em virtude de sua realimentação, o que significa que podem ser utilizadas em diversos sistemas variantes no tempo, como previsão de séries temporais, identificação e otimização de sistemas, controle de processos, etc

Reiterando a ideia de uma rede "realimentada", Vellasco (2007) enfatiza que os RNAs em circulação, também conhecidos como redes neurais artificiais com memória por possuírem realimentação, respondem aos estímulos de forma dinâmica. Em outras palavras, após a aplicação de uma nova entrada, o saída é calculada e então realimentada para modificar a entrada. Este procedimento é repetido inúmeras vezes, levando a mudanças sutis nas saídas que eventualmente podem se tornar persistentes. No entanto, mesmo com entradas consistentes, as redes neurais recorrentes nem sempre são estáveis.

#### 4 TREINAMENTO DA REDE NEURAL

Para Lefrançois (2001, p.6) temos que:

<sup>3</sup>Disponível em: [www.embarcados.com.br/redes-neurais-artificiais/](http://www.embarcados.com.br/redes-neurais-artificiais/). Acesso em 27 Novembro 2022

"Aprendizagem é definida como toda mudança relativamente permanente no potencial de comportamento humano, que resulta de experiência, mas não é causada por cansaço, maturação, drogas, lesões ou doenças. No sentido estrito, claro, a aprendizagem não é definida pelas mudanças reais ou potenciais no comportamento. Em vez disso a aprendizagem é o que acontece no organismo como resultado de experiências. As mudanças são simplesmente evidências de que a aprendizagem ocorreu"

Segundo Silva, Spatti e Flauzino (2010), o processo de treinamento de uma rede neural envolve a aplicação de etapas organizadas que são necessárias para a sincronização dos pesos sinápticos e limiares de seus neurônios, tendo como objetivo final a geração de soluções pelas saídas que são indicativas do sistema físico que estão mapeando.

A ideia de aprendizagem é crucial na caracterização de redes neurais artificiais. Isso é definido como uma alteração feita no peso da entrada de uma rede com a intenção de otimizar um mapa dos padrões de entrada e saída. Mesmo que seja possível ajustar manualmente a matriz  $W$ , sendo  $W$  uma matriz transposta, em alguns casos diretos para obter um mapa único, é preferível que esse processo seja realizado automaticamente. Em outras palavras, a rede deve ajustar automaticamente seus pesos à medida que o aprendizado avança por meio do uso de um algoritmo apropriado (SABBATINI, LINTON e GARCH, 1998).

Existem vários métodos de treinamento propostos para redes neurais artificiais, e estes são divididos em dois paradigmas principais: aprendizado não supervisionado e aprendizado supervisionado. Estes serão definidos e descritos nas seções a seguir.

#### 4.1 Treinamento sob supervisão

Esse aprendizado consiste em apresentar os sinais de entrada e saída para a rede. Uma parte considerável do tempo é dedicada a esse aprendizado. O treinamento em questão é considerado concluído quando a rede atinge um determinado nível de acerto. O conjunto de dados de treinamento deve ser suficientemente grande para conter as informações necessárias para que a rede aprenda subsequentemente os relacionamentos e padrões cruciais. De acordo com o resultado, o sistema precisa aprender como todos os fatores funcionam juntos antes de encontrar o ajuste mais adequado para os pesos em toda a rede.

Como resultado, os pesos são continuamente ajustados por meio de ações comparativas realizadas pelo próprio algoritmo de aprendizado, que monitora o desvio entre as respostas geradas pela rede e as desejadas, sendo a diferença utilizada no processo de ajuste. Quando tal defasagem estiver dentro dos limites aceitáveis, levando em conta os objetivos de generalização de soluções, a rede será considerada treinada.

O algoritmo de aprendizado supervisionado mais popular é chamado *Backpropagation* (mais informações sobre esse algoritmo serão fornecidas posteriormente). Existem três componentes na aprendizagem supervisionada que devem se destacar para a aprendizagem:

1. Ambiente (estacionário): As características estatísticas não mudam com o tempo. Desta forma, se o ambiente não for estacionário, as estatísticas mudam com o tempo. Ou, dito de outra forma, o ambiente estimula a rede neural apresentando padrões em sua cadência de entrada.
2. O professor (resposta pretendida): O processo de aprendizagem supervisionada funciona como se existisse um professor familiarizado com a resposta pretendida, o que significa que a sua organização interna é alterada em função destas normas e do resultado pretendido.
3. Algoritmo de aprendizado: ajuste os pesos para criar um mapa entrada-saída. As respostas futuras aos estímulos ambientais serão diferentes devido a mudanças estruturais internas. Existem muitos algoritmos para ajustar os parâmetros de uma rede neural artificial com base nessa ideia, incluindo *Adaline*, *Madaline* e retro propagação de erros.

## 4.2 Treinamento sem supervisão

Os resultados desejados estão ausentes quando um algoritmo de aprendizado não supervisionado é aplicado, em contraste com o aprendizado supervisionado. Como resultado, a rede interna deve se estabelecer em relação às individualidades existentes em relação aos componentes da coleção amostral total, identificando os grupos que compartilham semelhanças.

Isso significa que uma vez que a rede tenha sido ajustada aos parâmetros estatísticos dos dados de entrada, ela expande sua capacidade de criar representações internas para codificar as características de entrada e criar novas classes. Para este tipo de treino, pode ser aplicada a regra de aprendizagem competitiva.

Contudo a assimilação excessiva do modelo resulta na rede memorizando informações de treinamento, o que significa que um ambiente de treinamento exagerado faz com que a rede perca sua capacidade de generalizar a funcionalidade do modelo analógico de entrada. Como resultado, obtém-se um ótimo resultado dentro da amostra e um ajuste ruim fora da amostra. Além disso, empregar mais neurônios ocultos do que o necessário faz com que efeitos de entrada indesejáveis sejam contidos nos pesos sinápticos e faz com que a rede absorva ruído.

## 4.3 Perceptron da Camada Única

Nos anos de formação das redes neurais (1943-1958), alguns pesquisadores se sobressaíram devido as suas contribuições pioneiras:

- McCulloch e Pitts (1943) introduziram a ideia de redes neurais como máquinas de computação.
- Hebb (1949) propôs o perceptron como o primeiro modelo para aprendizagem dirigida pelo professor (ou seja, aprendizagem supervisionada).

Um Perceptron é uma forma mais simples de rede neural usada para classificar os chamados padrões linearmente separáveis (isto é, são padrões os quais podemos definir dois hiperplanos em um plano e assim poder classifica-los de acordo com os hiperplanos que definimos para nosso plano). Basicamente, consiste em um neurônio com pesos e bias sinápticos ajustáveis. Algoritmos para ajustar os parâmetros livres dessa rede neural apareceram pela primeira vez no processo de aprendizagem, desenvolvidos por Rosenblatt (1958-1962) como o modelo cerebral Perceptron. De fato, Rosenblatt provou que se os padrões (vetores) usados para terminar o Perceptron são retirados de duas classes linearmente separáveis, o algoritmo do Perceptron converge e coloca a superfície de decisão na forma de um hiperplano entre as duas classes.

A prova da convergência deste algoritmo é conhecida como teorema da convergência do Perceptron. Perceptrons construídos em torno de um único neurônio estão limitados a realizar a classificação de padrões com apenas duas classes (hipoteticamente). Ao estender a camada de saída do perceptron (computacional) para incluir vários neurônios, podemos classificar mais de duas classes de acordo. No entanto, as classes devem ser linearmente separáveis para que o perceptron funcione corretamente. O ponto importante é que, na medida em que consideramos a teoria básica dos perceptrons como classificadores de padrões, precisamos considerar apenas o caso de um único neurônio. É trivial estender essa teoria para o caso de mais um neurônio.

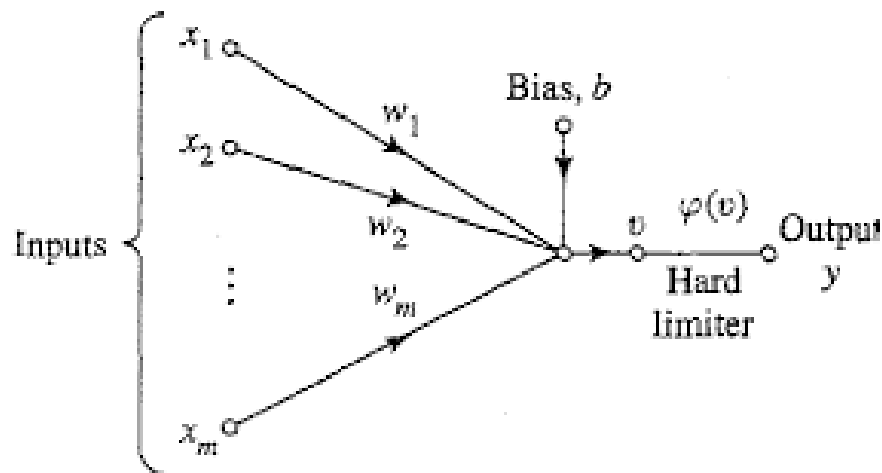
A título de estudos, podemos ter que os neurônios também formam a base dos filtros adaptativos, um bloco de construção funcional que fundamenta o tópico em constante expansão do processamento de sinais. O desenvolvimento da filtragem adaptativa deve-se em grande parte ao clássico artigo de Widrow e Hoff (1960), que criaram o chamado algoritmo dos mínimos quadrados médios (LMS), também conhecido como regra Delta. O algoritmo LMS é simples

de implementar, mas muito eficaz em termos de aplicação. Na verdade, é o carro-chefe da filtragem adaptativa linear, linear no sentido de que os neurônios operam em seu modo linear. Filtros adaptativos têm sido aplicados com sucesso em campos tão diversos como antenas, sistemas de comunicação, sistemas de controle, radar, sonar, sismologia e energia biomédica (Widrow e Stearns, 1985; Haykin, 1996).

## 5 ONDE A MATEMÁTICA ENTRA: O PERCEPTRON

O modelo de neurônio consiste em combinadores lineares e limitadores. Um nó adicional de um modelo neural calcula uma combinação linear das entradas aplicadas às suas sinapses e também incorpora vieses aplicados externamente. A soma resultante, o campo local induzido, é aplicada a um limitador. Correspondentemente, o neurônio produz uma saída igual a +1 se for positivo e -1 se for negativo dependendo da função de adaptação que estamos trabalhando, neste caso a função de ativação permite tais resultados.

Figura 6: Redes Neurais Princípios e Práticas



Fonte: HAYKIN (2001)

No modelo acima, que é do grafo de fluxo de sinais do Perceptron, temos que as entradas dos sinais para o grafo são dadas por  $x_1, x_2, \dots, x_m$ . Os pesos sinápticos do Perceptron estão representados por  $w_1, w_2, \dots, w_m$ . Logo após temos a aplicação do bias externo que é representado por  $b$ . Assim, para calcular o efeito das entradas juntamente com o bias, indicado por  $v$ , temos o seguinte cálculo

$$v = \sum_{i=1}^m w_i x_i + b \quad (1)$$

O objetivo desta rede neural, o Perceptron, é classificar todos os estímulos recebidos pela rede que são os  $x_1, x_2, \dots, x_m$  em dois classificadores,  $\xi_1$  ou  $\xi_2$ , de forma que para a classe  $\xi_1$  a saída do resultado do pceptron  $y$  seja +1 e para o classificador  $\xi_2$  a saída para o pceptron seja -1.

## 5.1 Convergência do Perceptron

Para que possamos melhor entender todo processo de convergência do Perceptron, vamos trabalhar com o grafo de fluxo estudado na seção anterior, sendo ele modificado. Teremos a presente demonstração que segue dos autores Carvalho, Braga & Ludermir (2000, p. 38-41) com algumas modificações, cabendo o crédito aos autores referenciados.

Considerando os sinais de entrada, para  $n$  estímulos,  $x_1(n), x_2(n), \dots, x_m(n)$  adicionados a +1 de forma a gerar um vetor de entrada com  $(m + 1)$  coordenadas, temos

$$x(n) = [+1, x_1(n), x_2(n), \dots, x_m(n)]^T \quad (2)$$

Sendo  $x(n)$  a matriz transposta. Definiremos os valores dos pesos do vetor  $x(n)$  como sendo

$$w(n) = [b(n), w_1(n), w_2(n), \dots, w_m(n)]^T \quad (3)$$

em que  $b(n)$  é o bias. Teremos assim o combinador linear gerando as saídas do sistema escrita na forma reduzida

$$v(n) = \sum_{i=1}^n w_i(n) \cdot x_i(n) = w^T(n) \cdot x(n) \quad (4)$$

onde  $w(n)$  representa o bias  $b(n)$  e ainda para um  $n$  fixo a seguinte equação  $w^T(n) \cdot x(n) = 0$  nos dá como resultado um espaço com coordenadas  $x_1, x_2, \dots, x_m$  onde este é um hiperplano com suas duas classes de entrada.

Trabalharemos agora as duas classes linearmente separáveis  $\xi_1$  e  $\xi_2$  já definidas previamente neste trabalho. Dado um conjunto de treinamento para um neurônio qualquer da camada de resposta com duas classes linearmente separáveis  $\xi_1$  e  $\xi_2$  digamos, para classificação, o conjunto de treinamento com  $n$  elementos  $\Gamma = [(x^i, 1)]_{i=1}^n$  pode ser dividido em dois subconjuntos

$$\begin{aligned} \Gamma_1 &= [(x^i, 1)]_{i=1}^{n_1} \in \xi_1 \\ \Gamma_2 &= [(x^i, 0)]_{i=1}^{n_2} \in \xi_2 \end{aligned}$$

As possíveis respostas que iremos obter passam a ser em relação a um problema binário, sendo que suas saídas irão ser com os números 1 e 0 de forma que tenhamos  $\Gamma = \Gamma_1 \cup \Gamma_2$  e  $n_1 + n_2 = n$ . Iremos admitir agora que exista solução para  $W_0$  sendo um vetor que adquire atais características neste momento:

$$\begin{aligned} W_0^T &= x^i \geq 0 \quad \forall x^i \in \xi_1 \\ W_0^T &= x^i < 0 \quad \forall x^i \in \xi_2 \end{aligned}$$

Portanto, o problema do treinamento se resume então a achar  $W_0$  que seja solução para a separação das duas classes  $\xi_1$  e  $\xi_2$  (Braga, Carvalho & Ludermir, 2000, p. 39). Essa seria a primeira forma de separação e classificação dos dados. Sendo a acurácia a precisão da operação e apuração dos valores, não obtendo 100% deste com a primeira rodada de pesos, durante o treinamento pode-se aplicar atualização dos pesos para minimizar os erros e encontrar a convergência. Os pesos serão atualizados da seguinte forma seguindo as condições:

$$\begin{aligned} w(t+1) &= w(t) + \lambda x(t) \quad \text{se } w^T(t)x(t) < 0 \quad \text{e } x(t) \in \xi_1 \\ w(t+1) &= w(t) - \lambda x(t) \quad \text{se } w^T(t)x(t) \geq 0 \quad \text{e } x(t) \in \xi_2 \end{aligned}$$

Logo o processo de atualização  $w(t+1) = w(t) + \lambda x(t)$  ocorre, no primeiro caso, quando o retorno da multiplicação das entradas com seus respectivos pesos for menor que 0; e, no segundo caso, atualiza caso a multiplicação retornar maior ou igual a 0 (em ambos os casos representando erro de classificação devido ao resultado esperado não ser atingido como esperado). Lembrando que a taxa de aprendizagem ( $\lambda$ ) tem influência no tempo de convergência, bom como a convergência pode intervir causando também a divergência. Agora, caso classifique incorretamente, digamos,  $k_1$  vetores da classe  $\xi_1$  e  $k_2$  vetores para  $\xi_2$  na qual  $k = k_1 + k_2$  apresenta as classificações incorretas no instante  $k$  como atualizar os pesos? Bem, com classes linearmente separáveis pode-se afirmar que existe uma solução  $w_0$ . Considerando as duas equações anteriores e  $\lambda = 1$ , podemos chegar na equação seguinte com  $w_0$ :

$$w_0^T w(t+1) = w_0^T \left( \sum_{i=1}^{t_1} x(t_1^i) \right) - \left( \sum_{i=1}^{t_2} x(t_2^i) \right)$$

$$w_0^T w(t+1) = w_0^T \left( \sum_{i=1}^t x(t) \right)$$

Note os sinais de adição e subtração na primeira e segunda equação anterior. Assim, facilmente chega-se na última formulação. Para uma solução fixa de  $w_0$  podemos buscar uma minimização:

$$\alpha = \min[w_0^T(x)t]$$

Com a desigualdade de Cauchy-Schwarz, pode-se concluir para os vetores  $w_0$  e  $w(t+1)$  que a desigualdade a seguir é válida

$$\|w_0^T\|^2 \|w(t+1)\|^2 \geq \|w_0^T, w(t+1)\|^2$$

Nos levando:

$$\|w_0^T\|^2 \|w(t+1)\|^2 \geq t^2 \alpha^2$$

$$\|w(t+1)\|^2 \geq \frac{t^2 \alpha^2}{\|w_0^T\|^2}$$

Assim, para ocorrer convergência é necessário que a norma  $w(t+1)$  seja limitada. Veremos mais a frente como se dará essa parte da demonstração. Com  $\lambda = 1$  pode-se escrever aquela equação de atualização dos pesos anterior como segue:

$$w(t+1) = w(t) + ex(t) \quad (5)$$

Obtendo a norma dos dois lados e elevando ambos os lados ao quadrado:

$$\|w(t+1)\|^2 = \|w(t) + ex(t)\|^2 \quad (6)$$

Expandido:

$$\begin{aligned} \|w(t+1)\|^2 &= [w_1(t)^2 + w_2(t)^2 + \dots + w_n(t)^2] \\ &\quad + e[x_1(t)^2 + x_2(t)^2 + \dots + x_n(t)^2] \\ &\quad + 2e[w_1(t)x_1(t) + w_2(t)x_2(t) + \dots + w_n(t)x_n(t)] \\ &= \|w(t)\|^2 + e\|x(t)\|^2 + 2ew^T(t)x(t) \end{aligned}$$

Com essa última formulação, obtendo erro igual a 1, isto é,  $e = 1$  dado  $d = 1$  e  $y = 0$ , lembre-se de  $e(t) = d(t) - y(t)$ , tem-se que

$$w^T(t)x(t) < 0. \quad (7)$$

O que implica

$$\|w(t)\|^2 + \|x\|^2 \geq \|w(t)\|^2 + e\|x(t)\|^2 + 2ew^T(t)x(t). \quad (8)$$

O que é verdade para  $e = -1$  também. Portanto:

$$\begin{aligned} \|w(t+1)\|^2 &\leq \|w(t)\|^2 + \|x(t)\|^2 \\ \|w(t+1)\|^2 - \|w(t)\|^2 &\leq \|x(t)\|^2 \end{aligned}$$

Somando os  $t$  primeiros termos da equação anterior:

$$\|w(t+1)\|^2 \leq \sum_{i=1}^n \|x(i)\|^2$$

Definindo:

$$\beta = \max_{i=1,2,\dots,t} \{\|x(i)\|^2\}$$

pode-se escrever outra condição de limite para o vetor  $\|w(t+1)\|$ :

$$\|w(t+1)\|^2 \leq \beta t$$

Portanto, tanto a formulação anterior quando  $\|w(t+1)\|^2 \geq \frac{t^2\alpha^2}{\|w_0\|}$  implicam na existência de um limite  $t$  para as operações de ajuste, já que:

$$0 \leq t \leq \frac{\beta\|w_0\|^2}{\alpha^2}$$

Assim,  $t = \frac{\beta\|w_0\|^2}{\alpha^2}$  limita o número de interações. Implicando sempre na convergência em tempo finito.

## 6 APLICAÇÃO

Após fazermos os estudos sobre a formas de treinamento do Perceptron, bem como sua convergência iremos trabalhar com exemplificação uma situação onde iremos aplicar os conceitos estudados. Considere inicialmente a separação e classificação de duas frutas distintas, laranjas e tangerinas, onde teremos a seguinte tabela:

Tabela 1: Exemplo: Aplicação do Treinamento do Perceptron

	$x_1$ (fosforo)	$x_2$ (acidez)	$x_3$ (cálcio)	CLASSIFICAÇÃO
AMOSTRA 1	0,1	0,4	0,7	1
AMOSTRA 2	0,2	0,3	0,1	1
AMOSTRA 3	0,6	0,9	0,8	-1
AMOSTRA 4	0,3	0,7	0,2	-1

Temos a tabela acima a qual as amostras 1 e 2 correspondem a fruta tangerina, e as amostras 3 e 4 correspondem a fruta laranja, sendo os resultados esperados após a conversão dos dados é 1 para tangerina e  $-1$  para laranja respectivamente. Teremos por interpretação matemática o seguinte somatório que iremos relacionar os dados de entradas, neste caso as amostra e os pesos envolvidos na operação.

$$\sum_{i=1}^3 x_i w_i + x_2 w_2 + x_3 w_3 + b$$

Usaremos como função de ativação a Degrau Bipolar que diz

$$\begin{cases} 1 & \text{se}(u > 0) \\ 0 & \text{se}(u = 0) \\ -1 & \text{se}(u < 0) \end{cases}$$

Através do estudo de uma aplicação devemos estar preparados para fazer o novo treinamento do sistema através da utilização de novos pesos sinápticos como segue

$$g(u) = \begin{cases} 1 & \text{se } u > 0 \\ 0 & \text{se } u = 0 \\ -1 & \text{se } u < 0 \end{cases} \quad \text{ou } g(u) = \begin{cases} 1 & \text{se } u \geq 0 \\ -1 & \text{se } u < 0 \end{cases}$$

Onde teremos as seguintes relações dos para gerar os novos pesos:

$$\begin{aligned} \Delta w_i &= \eta(Y_i - y_i)x_i \\ W_i &= w_i + \Delta w_i \end{aligned}$$

Onde temos  $\Delta w_i$  = delta peso;  $\eta$  = taxa de aprendizado  $Y_i$  = valor desejado  $y$  = saída do neurônio;  $x_i$  = dados de entrada;  $W_i$  = novo peso;  $w_i$  = peso antigo.

Vamos adotar  $\eta = 0,1$  como valor inicial para ser trabalhado. Iremos adotar  $w_1 = 0,1$ ;  $w_2 = 0,2$ ;  $w_3 = 0,3$  como valores iniciais para começarmos a estudar nossa aplicação adotando também uma *bias* = 0,1.

Chamaremos de Época 1 os teste iniciais com os pesos e o bias adotados previamente para este estudo para cada amostra da aplicação. seguinte tabela:

Tabela 2: Amostras - Época 1

ÉPOCA 1	$x_1$	$x_2$	$x_3$	VALOR DESEJADO	SAÍDA (y)
AMOSTRA 1	0,1	0,4	0,7	1	
AMOSTRA 2	0,2	0,3	0,1	1	
AMOSTRA 3	0,6	0,9	0,8	-1	
AMOSTRA 4	0,3	0,7	0,2	-1	

Note que os valores de saída nesta tabela acima não existe, pois so podemos afirmar o resultado quando aplicarmos os cálculos de convergência. Partindo para os cálculos em cada



amostra teremos:

$$\begin{aligned} AM(1) & : (0,1) \cdot (0,1) + (0,4) \cdot (0,2) + (0,7) \cdot (0,3) + 0,1 = 0,4 > 0 \\ AM(2) & : (0,2) \cdot (0,1) + (0,3) \cdot (0,2) + (0,1) \cdot (0,3) + 0,1 = 0,21 > 0 \\ AM(3) & : (0,6) \cdot (0,1) + (0,9) \cdot (0,2) + (0,8) \cdot (0,3) + 0,1 = 0,58 > 0 \end{aligned}$$

Tabela 3: Amostras - Época 1

ÉPOCA 1	$x_1$	$x_2$	$x_3$	VALOR DESEJADO	SAÍDA (y)
AMOSTRA 1	0,1	0,4	0,7	1	1
AMOSTRA 2	0,2	0,3	0,1	1	1
AMOSTRA 3	0,6	0,9	0,8	-1	1
AMOSTRA 4	0,3	0,7	0,2	-1	

Nesse momento notamos que o sistema presenteou uma falha pois o resultado esperado para a amostra (3) era  $-1$ , assim teremos que recalculamos os pesos do sistema até o momento em que este apresente o valor de saída do sistema seja o esperado. Assim ficamos com.

$$\begin{aligned} \Delta w_1 & = (0,1) \cdot (-1 - 1) \cdot (0,6) = -0,12 \\ \Delta w_2 & = (0,1) \cdot (-1 - 1) \cdot (0,9) = -0,18 \\ \Delta w_3 & = (0,1) \cdot (-1 - 1) \cdot (0,8) = -0,16 \\ W_1 & = 0,1 - 0,12 = -0,02 \\ W_2 & = 0,2 - 0,18 = 0,02 \\ W_3 & = 0,3 - 0,16 = 0,14 \end{aligned}$$

Assim, conseguimos encontrar novos pesos  $W_1$ ,  $W_2$  e  $W_3$ , aplicando na amostra (3) teremos:

$$AM(3) : (0,6) \cdot (-0,02) + (0,9) \cdot (0,02) + (0,8) \cdot (0,14) + 0,1 = 0,218 > 0$$

Vemos que ainda não alcançamos o resultado desejado, teremos que recalculamos os pesos. Então:

$$\begin{aligned} \Delta w_1 & = (0,1) \cdot (-1 + 1) \cdot (0,6) = 0 \\ \Delta w_2 & = (0,1) \cdot (-1 + 1) \cdot (0,9) = 0 \\ \Delta w_3 & = (0,1) \cdot (-1 + 1) \cdot (0,8) = 0 \\ W_1 & = (-0,12) + (-0,02) = -0,14 \\ W_2 & = (-0,18) + (0,02) = -0,16 \\ W_3 & = (-0,16) + (0,14) = -0,02 \end{aligned}$$

Aplicando os novos pesos na amostra (3) temos:

$$AM(3) : (0,6) \cdot (-0,14) + (0,9) \cdot (-0,16) + (0,8) \cdot (-0,02) + 0,1 = -0,144 < 0$$

Fazendo o teste para a amostra (4):

$$AM(4) : (0, 3) \cdot (-0, 14) + (0,7) \cdot (-0, 16) + (0, 8) \cdot (-0, 02) + 0, 1 = -0, 058 < 0$$

Então verificamos que com os novos pesos calculados  $W_1 = -0,14$ ,  $W_2 = -0,16$ , e  $W_3 = -0,02$  conseguimos obter os valores desejados para as amostras (3) e (4), logo, teremos que aplicar esses últimos pesos para toda a amostra iniciando a Época 2, ficamos da seguinte forma:

$$AM(1) : (0, 1) \cdot (-0, 14) + (0,4) \cdot (-0, 16) + (0, 7) \cdot (-0, 02) + 0, 1 = 0, 008 > 0$$

$$AM(2) : (0, 2) \cdot (-0, 14) + (0,3) \cdot (-0, 16) + (0, 1) \cdot (-0, 02) + 0, 1 = 0, 022 > 0$$

$$AM(3) : (0, 6) \cdot (-0, 14) + (0,9) \cdot (-0, 16) + (0, 8) \cdot (-0, 02) + 0, 1 = -0, 144 < 0$$

$$AM(4) : (0, 3) \cdot (-0, 14) + (0,7) \cdot (-0, 16) + (0, 2) \cdot (-0, 02) + 0, 1 = -0, 058 < 0$$

Assim alcançamos todos os valores desejados sobre prévia definição e ficamos com a seguinte tabela de resultados.

Tabela 4: Amostras - Época 2

ÉPOCA 2	$x_1$	$x_2$	$x_3$	VALOR DESEJADO	SAÍDA (y)
AMOSTRA 1	0,1	0,4	0,7	1	1
AMOSTRA 2	0,2	0,3	0,1	1	1
AMOSTRA 3	0,6	0,9	0,8	-1	-1
AMOSTRA 4	0,3	0,7	0,2	-1	-1

Onde chegamos nos resultados esperados através das adaptações dos pesos através do cálculo do  $\Delta w_1$

## 7 CONCLUSÃO

O desenvolvimento do presente estudo possibilitou uma análise das RNA's que podem ser aplicadas em diferentes situações nas mais diversas áreas.

Este trabalho teve como objetivo entender o funcionamento da rede neural Perceptron de uma camada utilizando previamente os conhecimentos biológicos da universidade celular neurônio, o qual dá base para o entendimento dos processos de aprendizado da máquina e conversão de dados no Perceptron. Com isso, pode-se perceber a necessidade de um estudo mais abrangente considerando o neurônio como ponto de partida.

Para se atingir uma compreensão dessa realidade da proposta do trabalho foram definidos dois objetivos específicos. O primeiro, de compreender a parte biológica a qual tivemos o estudo do sistema nervoso central do corpo humano, bem como as suas estruturas e a unidade central deste, o neurônio. Estudamos separadamente de forma sucinta cada uma das estruturas do sistema nervoso a fim de compreender o nosso segundo objetivo que é trabalhar com a rede neural Perceptron. Neste segundo objetivo, vimos como esta rede trabalha com padrão linear separável em duas classes pautado em um hiperplano estudado. Através da matemática aplicada, pudemos analisar como se dá a conversão de dados na máquina, adaptação de bias e resultados esperados através do treinamento proposto à máquina.

Após prévio estudo, podemos ver uma aplicação real na separação de dados de um experimento aleatório para a classificação de duas frutas, tangerina e laranja, através dos valores de

fósforo, acidez e cálcio das frutas. Nesta classificação podemos ver o treinamento da máquina a fim de nos dar um resultado esperado através do classificador binário.

Entendemos, assim, como funciona o treinamento do Perceptron, de modo a nos ofertar a resposta correta e repetir os valores em outras situações semelhantes. Por se tratar de um estudo introdutório, estudamos apenas o caso mais simples e o fizemos sem auxílio de programas de computador.

Em pesquisas futuras, pretendemos propor o estudo do Perceptron de múltiplas camadas a fim de compreender a fundo a abrangência dos artifícios e aplicações das Redes Neurais Artificiais, em especial a rede neural Perceptron e expandir os procedimentos para o uso de algoritmos em máquinas que é a finalidade do estudo de redes neurais artificiais.

## REFERÊNCIAS

- AMABIS, José M., MARTHO, Gilberto R.: **Biologia das Células**. 1. ed. São Paulo, Editora Modera, 2006
- BEAR, Mark F., CONNORS, Barry W., PARADISO, Michael A.: **Neurociências, desvendando o sistema nervoso**. Porto Alegre, EDITORA LTDA. 2017
- BURT, Alvin M.: **Neuroanatomia**. Rio de Janeiro, Guanabara Koogan S.A., 1995
- CARVALHO, André Carlos P. L. F. e BRAGA, Antonio de P. e LUDERMIR, Teresa B.: **Fundamentos de redes neurais artificiais**. Rio de Janeiro, 1998
- DUKES, Hond H.: **Fisiologia dos animais domésticos**. 11. ed. Rio de Janeiro: Guanabara Koogan, 1996.
- GAZZANIGA, Michael S.; IVRY, Richard B.; MANGUM, Goss R.: **Breve história da neurociência cognitiva**. In: Neurociência cognitiva: a biologia da mente. 2. ed. Porto Alegre: Artmed, 2006
- HAYKIN, Simon: **Redes Neurais: Princípios e prática**. Porto Alegre RS, 2001.
- LEFRANÇOIS, Guy R.: **Teorias da Aprendizagem**. São Paulo: Cengage Learning Petrópolis: Vozes, 2001
- MAGALHÃES, Lana: **Anatomia e Fisiologia Humana**. Blog Biologia .2013.
- MOREIRA, Carlos: **Neurônio**, Rev. Ciência Elem. São Paulo, 2008
- SABBATINI, M.; LINTON, O. GARCH.: **Implied Volatility Model** . International Journal of Forecasting, 1998.
- SILVA, Ivan N., SPATTI, Danilo H. e FLAUZINO, Rogério A.: **Redes neurais artificiais para engenharia e ciências aplicadas**. . São Paulo: Artliber Editora. 2010
- TRELEAVEN, Philip, PACHECO Mac A. C. e VELLASCO, Mark: **Architecture for Neural Network**. 9. ed. IEEE Micro, 1989.
- VELLASCO, Marly Maria Bernards R.: **Redes Neurais Artificiais**. Laboratório de Inteligência Computacional Aplicada, Rio de Janeiro, 2007

WIDROW, Bernard e STEARNS, Samuel D.: **Adaptive Signal Processing**. Tracey L Orbine. New Jersey. 1985

WIDROW, Bernard e HOFF, Marcian E.: **Adaptive Switching Circuits**, Ire Wescon Convention Record, 1960

## **AGRADECIMENTOS**

Em primeiro lugar, à Deus, que fez com que meus objetivos fossem alcançados, durante todos os meus anos de estudos, onde por muitas vezes foi minha força para continuar na caminhada.

Aos meus pais, que me incentivaram nos momentos difíceis e compreenderam a minha ausência enquanto eu me dedicava à realização das minhas pesquisas e estudos.

Aos professores, por todos os conselhos, pela ajuda e pela paciência com a qual guiaram o meu aprendizado.

Por fim, a todos aqueles que contribuíram, de alguma forma, para a realização deste trabalho