

Universidade Federal Fluminense

EVERALDO MARTINS ALEXANDRE

Diagnóstico Assistido por Inteligência
Computacional para Identificação de
Infecções

SANTO ANTÔNIO DE PÁDUA
2023

EVERALDO MARTINS ALEXANDRE

Diagnóstico Assistido por Inteligência Computacional para Identificação de Infecções

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Computacional em Ciência e Tecnologia da Universidade Federal Fluminense, como requisito parcial para obtenção do título de Mestre em Modelagem Computacional em Ciência e Tecnologia. Área de Concentração: Modelagem Computacional.

Orientador:

Cleber Almeida Corrêa Junior

Coorientadora:

Rosilene Abreu Portella Corrêa

Universidade Federal Fluminense

SANTO ANTÔNIO DE PÁDUA

2023

Ficha catalográfica automática - SDC/BEM
Gerada com informações fornecidas pelo autor

A381d Alexandre, Everaldo Martins
Diagnóstico Assistido por Inteligência Computacional para
Identificação de Infecções / Everaldo Martins Alexandre. -
2023.
68 f.: il.

Orientador: Cleber Almeida Corrêa Junior.
Coorientador: Rosilene Abreu Portella Corrêa.
Dissertação (mestrado)-Universidade Federal Fluminense,
Escola de Engenharia Industrial e Metalúrgica de Volta
Redonda, Volta Redonda, 2023.

1. Redes Neurais Artificiais. 2. Backpropagation. 3.
Hemograma. 4. Infecção Bacteriana. 5. Produção
intelectual. I. Junior, Cleber Almeida Corrêa, orientador.
II. Corrêa, Rosilene Abreu Portella, coorientadora. III.
Universidade Federal Fluminense. Escola de Engenharia
Industrial e Metalúrgica de Volta Redonda. IV. Título.

CDD - XXX

Diagnóstico Assistido por Inteligência Computacional para Identificação de Infecções

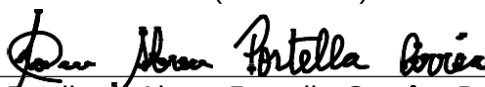
Everaldo Martins Alexandre

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Computacional em Ciência e Tecnologia da Universidade Federal Fluminense, como requisito parcial para obtenção do título de Mestre em Modelagem Computacional em Ciência e Tecnologia. Área de Concentração: Modelagem Computacional.

Aprovada por:



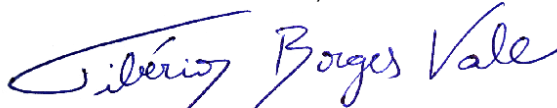
Prof. Cleber de Almeida Corrêa Junior, D.Sc. / MCCT-UFF
(Presidente)



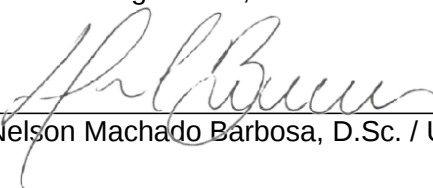
Prof^a. Rosilene Abreu Portella Corrêa, D.Sc. / MCCT-UFF



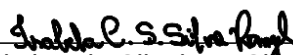
Prof. Ricardo Silveira Sousa, D.Sc. / MCCT-UFF



Prof. Tibério Borges Vale, D.Sc. / MCCT-UFF



Prof. Nelson Machado Barbosa, D.Sc. / UENF



Prof^a. Isabela Cristina da Silveira e Silva Rangel, D.Sc. /
FeMASS

Santo Antônio de Pádua, 14 de Março de 2023.

Agradecimentos

Agradeço primeiramente a Deus, por sempre ter me sustentado e ajudado ao longo de todo o caminho, e por mais difícil que seja a caminhada, Ele nunca me abandonou e sempre me deu forças para prosseguir.

Aos meus orientadores Cléber de Almeida Corrêa Junior e Rosilene Abreu Portella Corrêa, por terem me auxiliado durante a realização desta dissertação.

Ao professor Diomar Cesar Lobão em memória, por ter me orientado sobre todo processo de entrada e ter me incentivado a entrar no MCCT.

Aos meus pais Maria de Fátima Martins e Juscelino Oliveira, a minha esposa Taynarah Procópio, aos meus irmãos Alessandra Martins e José Geraldo Martins, por terem sempre me incentivado, mesmo durante os momentos de crise.

Aos professores Wesley Assis e Luís Alberto Duncan, por terem me ajudado sendo referências para que eu pudesse entrar no curso.

A todos professores e coordenação do MCCT, por todo apoio e ensinamentos durante o curso.

Aos profissionais da saúde Keyla Corrêa, Diogo Araújo e Marcelli Ferreira, por terem ajudado na avaliação dos hemogramas utilizados nesta dissertação.

A todos meus colegas de curso que também fazem parte desse momento, agradecimento especial a Louise Olmi e Márcio Geveg, que me ajudaram de alguma forma na dissertação.

As médicas que me ajudaram durante este momento que tenho enfrentado durante a realização deste trabalho.

E a todos demais que de alguma forma fizeram parte desse processo.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Resumo

O hemograma é um exame muito solicitado para se detectar várias doenças, tendo em vista que há possibilidade de auxiliar nas avaliações diagnósticas, e que existem profissionais que buscam auxílio para realizar a sua interpretação. Neste trabalho foi utilizado um método computacional que tem sido muito estudado e aplicado na medicina, que é a inteligência artificial. Sendo assim, dentro das inteligências artificiais, selecionou-se as Redes Neurais Artificiais (RNAs), estas que simulam o funcionamento dos neurônios humanos e são capazes de resolver problemas complexos. Inicialmente realizou-se uma breve introdução, buscando exemplos de aplicações de RNAs em análises médicas na literatura. A partir dos exames que estavam disponíveis e da possibilidade de auxílio aos profissionais da saúde. Buscou-se aplicar as RNAs para fazer a análise dos hemogramas, para verificar se estes apresentavam infecções bacterianas ou não. Para isso realizou-se uma revisão sobre infecções e hemograma completo. Também foi feita uma revisão sobre RNAs, apresentando com detalhes a rede que foi selecionada, que é a *Perceptron* multicamadas com algoritmo *backpropagation* e taxa de *momentum*. É apresentado seu funcionamento e como é realizado o treinamento desta rede. Para se fazer a aplicação deste método, solicitou-se que 3 profissionais da medicina analisassem os hemogramas quanto a presença de infecção bacteriana ou não. Com os resultados destes exames e observando que houve divergências entre as análises, foram criadas RNAs com as interpretações semelhantes. Por fim, analisou-se os resultados obtidos pelas RNAs e comparou-se com as avaliações que os profissionais da medicina divergiram.

Abstract

The complete blood count (CBC) is a very requested test to detect several diseases, considering that there is the possibility of helping in diagnostic evaluations and that there are professionals who seek help to perform its interpretation. In this work we used a computational method that has been widely studied and applied in medicine, which is artificial intelligence. Artificial Neural Networks (ANNs), which simulate the functioning of human neurons and can solve complex problems, were selected among the artificial intelligences. Initially, a brief introduction was made looking for examples of ANN applications in medical analysis in the literature. Based on the available tests and the possibility of helping health professionals. RNNs were applied to analyze blood counts to verify if they showed bacterial infections or not. For this, a review about infections and complete blood count was performed. A review about ANNs was also done, presenting in detail the network that was selected, which is the multilayer Perceptron with backpropagation algorithm and momentum rate. Its operation is presented and how the training of this network is performed. To apply this method, it was requested that 3 medical professionals analyze the hemograms as to the presence of bacterial infection or not. With the results of these exams and observing that there were divergences between the analyses, ANNs were created with similar interpretations. Finally, the results obtained by the ANNs were analyzed and compared with the evaluations and that the medical professionals diverged.

Palavras-chaves

1. Redes Neurais Artificiais
2. Infecções Bacterianas
3. Doenças
4. Hemograma
5. *Backpropagation*

Lista de Figuras

Figura 2.1: Neutrófilo Bastonado.	20
Figura 2.2: Neutrófilo Segmentado.	21
Figura 3.1: Conexões sinápticas entre neurônios.	23
Figura 3.2: Neurônio Artificial.	24
Figura 3.3: Função de ativação degrau.	26
Figura 3.4: Função de ativação degrau bipolar.	26
Figura 3.5: Função de ativação rampa simétrica.	27
Figura 3.6: Função de ativação logística.	28
Figura 3.7: Função de ativação tangente hiperbólica.	29
Figura 3.8: Função de ativação gaussiana.	30
Figura 3.9: Função de ativação linear.	30
Figura 3.10: Arquitetura feedforward de camada simples	31
Figura 3.11: Arquitetura feedforward de camadas múltiplas.	32
Figura 3.12: Arquitetura recorrente ou realimentada.	33
Figura 3.13: Arquitetura em estrutura reticulada.	33
Figura 3.14: Exemplo de rede Perceptron multicamadas.	35
Figura 3.15: As duas fases de treinamento da rede PMC.	37
Figura 3.16: Notação para derivação do algoritmo backpropagation.	38
Figura 3.17: Configuração de neurônio usado na derivação do algoritmo backpropagation	39
Figura 4.1: Desenho esquemático da RNA usada	53
Figura 4.2: Gráfico de distribuição do sexo dos pacientes	54
Figura 4.3: Treinamento da RNA com 100 neurônios.	55
Figura 4.4: Gráfico de valor esperado x RNA com 100 neurônios	56
Figura 4.5: Treinamento da RNA com 200 neurônios	57
Figura 4.6: Gráfico de valor esperado x RNA com 200 neurônios	58
Figura 4.7: Treinamento da RNA com 500 neurônios	59
Figura 4.8: Gráfico de valor esperado x RNA para 500 neurônios	60

Lista de Tabelas

Tabela 1. Valores de Referência do Eritrograma.....	19
Tabela 2. Valores de Referência do Leucograma.....	21
Tabela 3. Valor Esperado x RNA com 100 neurônios.....	56
Tabela 4. Valor Esperado x RNA com 200 neurônios.....	58
Tabela 5. Valor Esperado x RNA com 500 neurônios.....	60
Tabela 6. RNA com 100 neurônios x Profissionais da medicina	61
Tabela 7. RNA com 200 neurônios x Profissionais da medicina	62
Tabela 8. RNA com 500 neurônios x Profissionais da medicina	63

Sumário

1	Introdução	12
1.1	Objetivos	14
1.1.1	Objetivo Geral	14
1.1.2	Objetivos específicos	15
1.2	Organização do Trabalho	15
2	Hemograma Completo e Infecção	16
2.1	Infecção Bacteriana	17
2.2	Eritrograma	17
2.3	Leucograma	19
3	Redes Neurais Artificiais	22
3.1	Funções de Ativação	25
3.2	Arquitetura <i>Feedforward</i> de Camadas Simples	31
3.3	Arquitetura <i>Feedforward</i> de Múltiplas Camadas	31
3.4	Arquitetura Recorrente ou Realimentada	32
3.5	Arquitetura em Estrutura Reticulada	33
3.6	Tipos de Treinamento para Rede Neural Artificial	34
3.7	Redes <i>Perceptron</i> Multicamadas	34
3.7.1	Funcionamento da rede <i>Perceptron</i> multicamadas	35
3.7.2	Processo de treinamento do <i>Perceptron</i> multicamadas	36
3.7.3	O Algoritmo <i>Backpropagation</i>	38
4	Resultados e discussão	52
5	Conclusões e trabalhos futuros	64
	Referências	66

Capítulo 1

1 Introdução

A saúde é um dos maiores bens que o ser humano possui, e o seu valor é maior do que qualquer dinheiro do mundo para muitas pessoas. Sendo assim, é recomendado consultar-se regularmente com um médico, para que se possa evitar possíveis doenças e ter, se for o caso, um possível diagnóstico precoce destas. O hemograma é o exame mais solicitado para detecção das mais variadas doenças e segundo Rosenfeld (2012), o hemograma pode ser realizado em até um minuto.

Muitos médicos estão se formando e os recém-formados carregam muita dúvida consigo, e até mesmo vários médicos experientes enfrentam esse problema, como pôde ser notado pelos comentários realizados por muitos profissionais da saúde que participaram do curso de Interpretação de Exames Laboratoriais, palestrado pelo professor Rodrigo de Carvalho (CARVALHO, 2021). Para ter maior segurança na análise desses exames, muitos médicos utilizam-se da pesquisa online para auxiliá-los nesse trabalho.

A medicina tem evoluído muito rapidamente com o auxílio da tecnologia de forma geral, e conforme os métodos computacionais avançam, eles se tornam grandes ferramentas para a melhoria do diagnóstico de doenças.

Dentre os vários avanços computacionais que têm sido estudados atualmente, um que tem sido muito utilizado nas diversas áreas da ciência e na área médica é a inteligência artificial. A inteligência artificial é realizada através de computadores que operam por meio de algoritmos e propõem soluções para problemas médicos (LOBO, 2017).

Dentre os diversos tipos de inteligências artificiais existem as Redes Neurais Artificiais (RNAs), essas que se assemelham ao funcionamento das Redes Neurais biológicas humanas e foi o método utilizado neste trabalho. As RNAs são um exemplo de método computacional muito estudado e aplicado para diagnóstico de doenças (SIQUEIRA-BATISTA et al., 2014).

Neste trabalho buscou-se criar uma RNA que pudesse auxiliar na análise de hemogramas para se detectar infecções bacterianas. Pois essas infecções podem ser

tratadas com antibióticos, e tendo em vista que existe casos em que estes remédios são receitados de forma discriminada. Por meio da RNA poderia auxiliar no processo de tomada de decisão dos médicos e assim diminuir o excesso de uso de antibióticos. Além disso poderia ajudar aqueles profissionais que se sentem ainda inseguros a terem uma segunda opinião, essa que teria como base para o treinamento da RNA, as avaliações realizadas por uma junta médica.

A fim de mostrar que as Redes Neurais possuem grande aplicação, a seguir são apresentados alguns exemplos de utilização de Redes Neurais, auxiliando na resolução de diversos problemas.

No ramo da medicina pode-se citar o trabalho realizado por Helegda et al. (2015), que realizou uma aplicação que se assemelha ao que foi feito neste trabalho, porém eles trabalham com dados para se detectar doenças cardiovasculares. Eles utilizaram cerca de 80 amostras, sendo que metade delas apresentava a doença, e a outra metade não apresentava. Para o funcionamento da rede usou-se 15 parâmetros para entrada e 1 parâmetro de saída. Ao final dos testes a Rede Neural teve um excelente desempenho, acertando mais de 90 % dos exames testados na rede.

De acordo com Goulart (2006), conjuntivites alérgicas (CA) são muito frequentes no Brasil, sendo assim, ele buscou uma forma de auxiliar na verificação dos pacientes, para se identificar os que possuem ou não o problema. Para este estudo utilizou-se 48 crianças que eram acompanhadas por médicos, que possuíam a doença e se queixavam, pois tinham desconfortos oculares e outras 54 crianças que não apresentavam nenhum problema e nem se queixavam. As crianças que participaram desse estudo possuíam de 3 a 14 anos. Após isso foi feito um questionário com 15 perguntas, para que as crianças pudessem responder sobre os desconfortos oculares enfrentados, em sua maioria com respostas positivas ou negativas. De todas as 15 perguntas feitas, apenas 7 foram usadas como dados de entrada para a criação da Rede Neural. Foram usados 61 questionários para o treinamento da Rede Neural e os outros 41 utilizados para o teste dessa rede, todos estes escolhidos aleatoriamente. Após todos os testes a Rede Neural Artificial obteve 100 % de acerto dos questionários (GOULART, 2006).

De acordo com Veronezi et al. (2011), osteoartrite (OA) é um transtorno articular que causa dor e limitações funcionais. A (OA) é um problema muito comum que afeta as articulações, seja no Brasil ou no mundo. Para ajudar na avaliação dos pacientes, Veronezi et al. (2011) utilizou 206 radiografias da coluna lombar, obtidas

através de um serviço de radiologia. Após coletar todas essas imagens, Veronezi et al. (2011) treinou uma Rede Neural com o objetivo de analisar as imagens e dizer se estas apresentam ou não osteoartrite. Para o treinamento dessa rede foram utilizadas 68 radiografias, para teste usou-se 68 radiografias e as outras 70 foram usadas para validação. Depois de ter realizado todo o processo, a RNA obteve 62,85 % de acurácia, como descrito pelo autor foi muito abaixo da média de 85 % encontrada na literatura para utilização de RNA, possivelmente devido a distribuição das amostras para treinamento e testes, não terem sido bem distribuídas, ou a escolha aleatória pode não ter favorecido o treinamento da RNA. Porém ainda é um bom resultado para aplicação de RNA (VERONEZI et al., 2011).

Segundo Costa, Motta e Nogueira (2010), as doenças meningocócicas levam um indivíduo rapidamente a morte. Por isso devem ser descobertas precocemente, para que possam ser tratadas e a saúde venha a ser restaurada. Sendo assim, o trabalho de Costa, Motta e Nogueira (2010) propôs a criação de uma Rede Neural Artificial para auxiliar no diagnóstico dessas doenças. Para criar esta RNA foram utilizados 125 exames, sendo eles usados para treinamento e teste, e após esse processo, a rede iria classificar os exames em 7 diferentes tipos de doenças meningocócicas. Finalizados todos os treinamentos e testes, a Rede Neural obteve uma ótima eficiência na classificação, obtendo os seguintes acertos: meningocemia 80 %, meningocócica 82 %, tuberculosa 62,5 %, outra bactéria 70 %, asséptica 100 %, outra etiologia 100 %, pneumocócica 100 % (COSTA; MOTTA; NOGUEIRA, 2010).

A partir dos exemplos citados, pode-se perceber que as Redes Neurais têm várias aplicações e podem ser amplamente utilizadas na medicina. Por este motivo foi o método escolhido para ser aplicado nesse trabalho, na análise e interpretação dos exames e verificação se esses apresentam infecções bacterianas ou não.

1.1 Objetivos

1.1.1 Objetivo Geral

O presente trabalho tem por objetivo geral usar Redes Neurais Artificiais no processo de análise de hemogramas, a fim de detectar se o exame apresenta uma infecção bacteriana ou não.

1.1.2 Objetivos específicos

Os objetivos específicos deste trabalho são:

- Realizar uma revisão da literatura sobre hemogramas e os parâmetros importantes para detectar infecções bacterianas.
- Fazer um estudo bibliográfico sobre Redes Neurais Artificiais.
- Realizar o treinamento de uma Rede Neural que avalie através do hemograma completo, se o paciente apresenta infecção bacteriana ou não.
- Testar a Rede Neural criada.
- Comparar os resultados obtidos pelas Redes Neurais com as avaliações médicas dos hemogramas, a fim de avaliar sua eficiência.

1.2 Organização do Trabalho

A dissertação está organizada em 5 Capítulos. Neste Capítulo é apresentada a motivação de se ter escolhido esse tema, uma breve revisão bibliográfica sobre alguns exemplos de aplicação de RNAs na medicina e quais os objetivos a serem alcançados pelo trabalho.

No Capítulo 2 apresenta-se uma explicação sobre infecções e hemogramas, além de explicações sobre os parâmetros utilizados para interpretação desse.

O Capítulo 3 traz-se uma explicação sobre Redes Neurais, os tipos de Redes Neurais existentes e uma breve explicação sobre elas, além de uma explicação detalhada da rede neural artificial que será utilizada no trabalho, com as formulações matemáticas.

No Capítulo 4 apresenta-se os resultados obtidos pelo treinamento da Rede Neural Artificial utilizada, e é feita uma breve discussão sobre esses.

Finalmente no Capítulo 5, faz-se as devidas conclusões do trabalho e sugestões para trabalhos futuros.

Capítulo 2

2 Hemograma Completo e Infecção

Segundo Santos (2021), infecção é “a penetração ou desenvolvimento de um agente infeccioso no organismo do homem ou de outro animal”. Os agentes infecciosos que podem causar essas doenças infecciosas são divididos principalmente em 4 tipos: vírus, bactérias, fungos e protozoários. Para este trabalho serão estudadas infecções causadas por bactérias.

A febre é um importante aliado na luta contra infecções, pois a maioria das bactérias e vírus causadores das infecções se proliferam melhor abaixo de 37° C. Sendo assim aumentar alguns graus da temperatura corporal, auxilia no combate as infecções e ajudam o corpo a vencê-las. A febre faz com que o sistema imunológico produza mais glóbulos brancos, anticorpos e outros agentes que ajudam a enfrentar as infecções (FALCON DÍAZ, 2009).

O sangue humano é constituído principalmente por eritrócitos (hemácias), séries brancas (leucócitos) e plaquetas (DE ALMEIDA; DE MELO; CÂMARA DE ALMEIDA, 2012). Para se detectar infecções através de um hemograma é preciso saber quais parâmetros compõe um hemograma e quais são mais importantes para este. Geralmente quando um médico solicita um hemograma, ele solicita o leucograma e o eritograma, dados estes que serão utilizados para alimentar a rede neural, que será utilizada nesta pesquisa.

No hemograma as infecções causam alterações primeiramente na série branca, aumentando ou diminuindo seus níveis, de acordo com o tipo de agente infeccioso ou o estado imunológico do paciente (RAMOS; SOUZA, 2016).

Segundo Failace e Fernandes (2015), o hemograma é o exame mais pedido pelos médicos e sua avaliação quantitativa é feita de forma automatizada através de um software. Por outro lado, a avaliação qualitativa ainda não é feita com perfeição usando a forma computadorizada. De acordo com Bandeira, Magalhães e de Aquino (2014), atualmente os laboratórios fazem seus exames de hemograma de forma automatizada, pois assim os resultados obtidos são mais sensíveis e precisos. Existem diversos aparelhos para realizar essa análise no mercado.

2.1 Infecção Bacteriana

Segundo Salomão (2017), as doenças infecciosas resultam da relação entre seres vivos, o patógeno e o hospedeiro. No intestino de uma pessoa de 70 kg há cerca de 100 trilhões de micróbios. As bactérias vivem de forma simbiótica no organismo humano. Em torno de 1 % das bactérias são patogênicas. Até as bactérias que fazem parte da microbiota humana podem se transformar em patogênicas, de acordo com as reações que o corpo venha a sofrer. O ser humano também pode ser contaminado por bactérias através do ar, solo e água que estejam contaminados.

De acordo com Santos (2014), as bactérias são parte natural do ambiente em que vivemos. Fazem parte do revestimento da pele e das mucosas, e cobrem o trato intestinal do animal e humano. Grande parte das bactérias são inofensivas, e podem até auxiliar o seu hospedeiro, evitando que bactérias nocivas se proliferem. Como as bactérias são geradas rapidamente em minutos ou horas, elas podem reagir as mudanças do ambiente rapidamente.

Segundo Salomão (2017), nas doenças bacterianas ocorre o aumento da quantidade de leucócitos que é chamada de leucocitose, causando também a neutrofilia (aumento dos neutrófilos) que pode ser relativa ou absoluta.

Doenças bacterianas são frequentemente acompanhadas pelo aumento das formas mais jovens de granulócitos, tanto em número absoluto quanto percentual, que é denominado de desvio à esquerda. Há também casos em que se tem leucopenia absoluta, com neutropenia, normalmente nos casos em que as infecções são causadas por bactérias Gram-positivas, em casos de tuberculose, brucelose, entre outros, nessas situações, elas podem apresentar ou não o desvio à esquerda. Os neutrófilos são compostos por bastões e segmentados (SALOMÃO, 2017).

2.2 Eritrograma

Segundo Brazão-Silva (2019), o eritrograma apresenta a contagem de hemácias ou eritrócitos (CE), hematócrito (Ht), dosagem de hemoglobina (Hb) e os índices hematimétricos, estes últimos que são divididos em: volume corpuscular médio (VCM), hemoglobina corpuscular média (HCM), concentração da hemoglobina

corpuscular média (CHCM) e *red blood cell distribution Width* (RDW) que é a amplitude de distribuição dos glóbulos vermelhos.

As hemácias são unidades morfológicas que estão no sangue na quantidade de cerca de $4,5$ a $6,5 \times 10^6/\text{mm}^3$, em condições normais. Elas são constituídas essencialmente por globulina e hemoglobina. hemácias normais apresentam-se na forma de discos circulares homogêneos, bicôncavos e sem núcleo (DE ALMEIDA; DE MELO; CÂMARA DE ALMEIDA, 2012).

O hematócrito (Ht) representa a proporção entre a parte sólida e líquida do sangue. A hemoglobina é a principal proteína encontrada no sangue e ela realiza as trocas gasosas e oxidação residual (BRAZÃO-SILVA, 2019). De acordo com Failace e Fernandes (2015), o hematócrito é o volume da massa eritróide de uma amostra de sangue, expresso em porcentagem ou fração decimal do volume desta. Os valores de referência apresentados são homens $47 \pm 7 \%$ e mulheres $42 \pm 6\%$.

Carvalho et. al. (2016), diz que “o volume corpuscular médio (VCM) é o índice hematimétrico absoluto pelo volume globular e pela contagem eritrocitária.” O VCM analisa o tamanho e/ou volume médio dos eritrócitos. Segundo Failace e Fernandes (2015), a ideia de VCM foi criada e difundida por Wintrobe (1934). O VCM é calculado através da divisão do hematócrito pela contagem de eritrócitos. Eles são importantes para se fazer a classificação das anemias em micro, normo e macrocíticas. Os valores de referência do VCM, são tanto para homem quanto para mulher em torno de $88-90 \pm 7 \text{ fL}$, quando o VCM é maior que 98 fL há uma macro citose e valor do VCM for menor que 82 fL tem-se uma microcitose. Com os avanços das tecnologias, nos exames computadorizados podem já vir impressos após o VCM, se há uma macro citose ou micro citose.

Segundo Brazão-Silva (2019), o HCM avalia a quantidade média em massa de hemoglobina por hemácias (variando de 24 a 33 pg). De acordo com Failace e Fernandes (2015), o cálculo do HCM é feito de forma computadorizada, por meio da divisão da dosagem de hemoglobina pelo número de eritrócitos. O valor do HCM é exato, pois tanto a hemoglobina quanto os eritrócitos apresentam valores exatos

De acordo com Brazão-Silva (2019), o CHCM representa a porcentagem de hemoglobina dentro das hemácias (variando de 32 a 36%).

Apesar da literatura dar padrões para os parâmetros hematológicos, normalmente cada laboratório segue um determinado padrão. Com a finalidade de se observar os valores normais de cada um dos parâmetros de um hemograma, de

acordo com o laboratório do hospital de Santo Antônio de Pádua que forneceu os exames, sendo estes os padrões seguidos. Criou-se então a Tabela 1 com os valores de referência de cada um desses dados apresentados nos exames na série vermelha.

Tabela 1. Valores de Referência do Eritrograma.

	Valores Normais			
	Homens		Mulheres	
	Mínimo	Máximo	Mínimo	Máximo
Hemácias	4,50 x10 ⁶ /mm ³	6,00x10 ⁶ /mm ³	4,00 x10 ⁶ /mm ³	5,60 x10 ⁶ /mm ³
Hemoglobina	14,00 g/dL	17,00 g/dL	8,50 g/dL	12,00 g/dL
Hematócrito	37 %	50 %	36 %	47 %
V.C.M.	80 fL	94 fL	80 fL	94 fL
H.C.M.	27 pg	32 pg	27 pg	32 pg
C.H.C.M.	32 g/dL	38 g/dL	32 g/dL	38 g/dL

Fonte: Autoria própria, 2023.

2.3 Leucograma

O leucograma é composto pelos leucócitos, que se dividem em basófilos, eosinófilos, mielócitos, jovens, bastões, segmentados, linfócitos e monócitos. Cada um deles são apresentados em porcentagem, mostrando o quanto representam na composição dos leucócitos e a quantidade total de cada um.

Os leucócitos são unidades móveis do sistema mononuclear fagocitário e flutuam livremente pelo sangue, eles possuem a capacidade de fagocitar ou eliminar micro-organismos estranhos. Os leucócitos são muito importantes, pois eles podem se transportar até as zonas onde encontram-se as inflamações, proporcionando uma defesa rápida e forte contra os agentes infecciosos (FALCON DÍAZ, 2009)

De acordo com Failace e Fernandes (2015), cada pessoa possui valores normais de contagem de leucócitos, sendo assim, faz-se necessário guardar exames que forem realizados, para que se possa comparar com exames feitos no futuro. Quando a quantidade de leucócitos está acima dos valores normais, tem-se uma leucocitose, e quando essa mesma quantidade está abaixo dos valores considerados normais tem-se a chamada leucopenia.

O aumento ou diminuição do número de leucócitos sozinho não é suficiente para se interpretar o que está acontecendo com o paciente, por isso utiliza-se da

fórmula leucocitária que começou com Ehrlich (1879). Esta fórmula resume-se basicamente em observar a porcentagem da composição dos leucócitos, como por exemplo, quando se tem uma leucocitose e observa-se que ao mesmo tempo o número de mielócitos está anormal ou o número de eosinófilos, tendo assim uma interpretação do que está ocorrendo com o paciente (FAILACE; FERNANDES, 2015).

Conforme escreveu Failace e Fernandes (2015), eosinófilos, basófilos e monócitos são células mielóides e, da mesma forma que os neutrófilos, procedem de predecessor comum. Eles possuem cada um, uma função diferente, esta que é percebida através das alterações no hemograma, gerando de acordo com o aumento ou a diminuição de cada um deles, várias doenças.

Segundo Chauffaille (2010), o eosinófilo é uma célula com cerca de 8 a 15 μm de diâmetro, possui núcleo normalmente bilobado e que apresenta grânulos intracitoplasmáticos com alta afinidade por eosina. Os eosinófilos representam 3% aproximadamente das células que compõe a medula óssea de indivíduos normais.

Os neutrófilos segmentados e bastonados são células que possuem maturação e funções semelhantes e são apresentados nas Figuras 2.1 e 2.2.

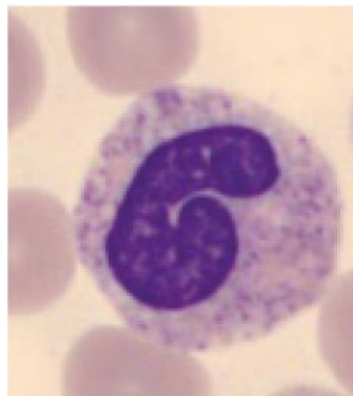


Figura 2.1: Neutrófilo Bastonado.

Fonte: Failace e Fernandes, 2015.

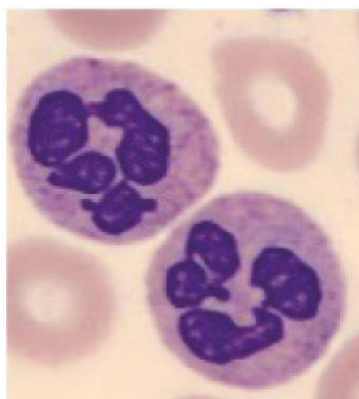


Figura 2.2: Neutrófilo Segmentado.

Fonte: Failace e Fernandes, 2015.

Com a finalidade de se auxiliar melhor na análise dos dados da série branca, criou-se a Tabela 2, a partir dos valores normais estabelecidos pelo laboratório do hospital de Santo Antônio de Pádua, onde conseguiu-se os exames.

Tabela 2. Valores de Referência do Leucograma.

	Valores Normais			
	Mínimo (%)	Máximo (%)	Mínimo (mm ³)	Máximo (mm ³)
Leucócitos			5000	10000
Basófilos	0	1	0	100
Eosinófilos	1	4	100	400
Mielócitos	0	0	0	0
Jovens	0	1	0	100
Bastões	3	5	150	500
Segmentados	50	67	2500	6700
Linfócitos	21	35	1050	3500
Monócitos	4	8	200	800

Fonte: Autoria própria, 2023.

Capítulo 3

3 Redes Neurais Artificiais

Neste trabalho utilizou-se as Redes Neurais Artificiais (RNAs), pois unem as melhores capacidades dos computadores com as do cérebro humano. O cérebro humano processa as informações de forma não linear e em paralelo, dessa forma, ele é mais rápido que o computador no reconhecimento de padrões, por exemplo (RAMOS, 2002).

Dentro do cérebro humano o processamento de informações é feito por elementos processadores biológicos operando em paralelo. O neurônio é a célula elementar do sistema nervoso cerebral. Ele é responsável por conduzir impulsos (estímulos elétricos advindos de reações físico-químicas) (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

De acordo com Silva, Spatti e Flauzino (2010), como representado na Figura 3.1, o neurônio pode ser dividido em 3 partes fundamentais. A primeira é chamada de dendritos, que são formados por muitos finos prolongamentos, formando a árvore dendrital. Eles possuem a função de receber continuamente os estímulos provenientes de outros neurônios, ou que vem do meio externo onde esses podem estar em contato (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

A segunda chama-se Corpo Celular ou Soma, e é responsável pelo processamento das informações provenientes dos dendritos com a finalidade de produzir um potencial de ativação, que mostrará se o neurônio irá lançar um impulso elétrico ao longo do axônio.

A terceira é denominada Axônio, que é formado por um longo filamento. O axônio tem por função fazer a condução dos impulsos elétricos para outros neurônios. A sua extremidade é formada por ramificações chamadas terminações sinápticas (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

Através das sinapses é feito a conexão do axônio de um neurônio e os dendritos de outros neurônios, por meio de impulsos elétricos. Como apresentado na Figura 3.1. Não existe contato físico entre os neurônios na junção (fenda) sináptica. São os elementos neurotransmissores liberados que fazem a ponderação dos

impulsos elétricos entre os neurônios (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

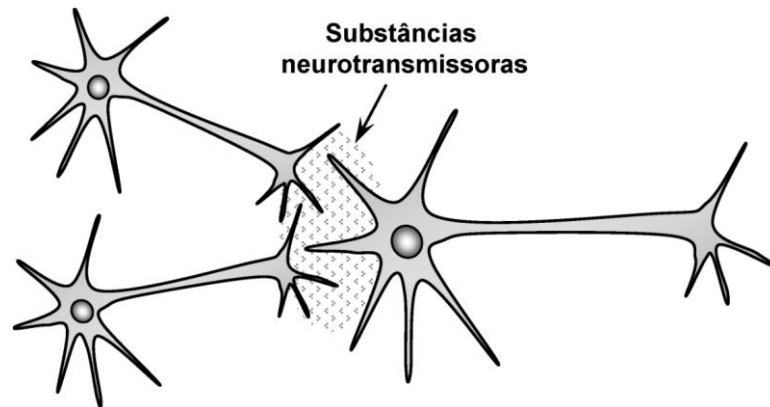


Figura 3.1: Conexões sinápticas entre neurônios.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

As RNAs foram criadas a partir de modelos do sistema nervoso humano. Os neurônios artificiais são modelos simplificados dos neurônios biológicos. Esses modelos têm por inspiração a análise da geração e propagação dos impulsos elétricos, através da membrana celular dos neurônios (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

McCulloch e Pitts (1943), criaram o modelo mais usado para definir as RNAs. Um exemplo de como o neurônio artificial pode ser modelado está representado na Figura 3.2. Os sinais de entradas provenientes do meio externo são representados por $\{x_1, x_2, \dots, x_n\}$, que correspondem aos impulsos elétricos externos que os dendritos recebem. Os pesos sinápticos são representados por $\{w_1, w_2, \dots, w_n\}$ e são análogos as ponderações que as junções sinápticas exercem. A importância de cada uma das entradas $\{x_i\}$ dos neurônios é determinada através da multiplicação por seus pesos sinápticos $\{w_i\}$, resultando na ponderação de toda informação de entrada. Dessa forma, a saída do corpo celular artificial representada por $\{u\}$, é a soma ponderada das entradas (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

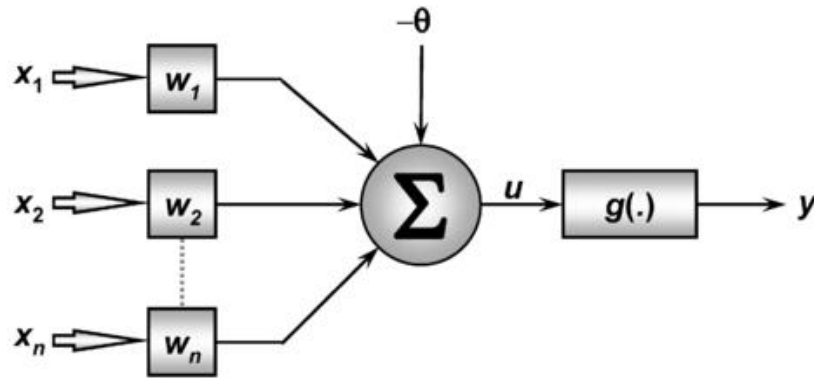


Figura 3.2: Neurônio Artificial.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

A partir da Figura 3.2, de acordo com Silva, Spatti e Flauzino, 2010 e Pereira, 2018, pode-se observar que o neurônio artificial é formado de 7 elementos básicos, que são:

- **Sinais de entradas** $\{x_1, x_2, \dots, x_n\}$: são os sinais ou medidas oriundas do meio externo, e representam os valores que as variáveis de determinada aplicação assumem.
- **Pesos Sinápticos** $\{w_1, w_2, \dots, w_n\}$: são os valores usados para ponderar cada uma das variáveis de entrada da rede, possibilitando que as suas importâncias sejam medidas em relação à funcionalidade do neurônio correspondente.
- **Combinador Linear** $\{\Sigma\}$: é responsável por somar todos os sinais de entrada ponderados pelos seus pesos sinápticos correspondentes, com o objetivo de criar o potencial de ativação.
- **Limiar de Ativação** $\{\theta\}$: é uma variável que determina qual o nível adequado para que o resultado realizado pelo combinador linear consiga produzir um valor de disparo rumo à saída do neurônio.
- **Potencial de Ativação** $\{u\}$: resulta da diferença entre o valor gerado pela combinação linear e o limiar de ativação.
- **Função de Ativação** $\{g\}$: tem a função de limitar a saída do neurônio, respeitando um intervalo de valores aceitáveis que a sua própria imagem funcional possa receber.
- **Sinal de Saída** $\{y\}$: é o valor final que o neurônio produz com relação a um determinado grupo de sinais de entrada, outros neurônios que estão interligados em sequência, também podem utilizá-lo.

O funcionamento do neurônio artificial da Figura 3.2 proposto por McCulloch e Pitts (1943) expresso em notação matemática, pode ser representado conforme as Equações (3.1) e (3.2).

$$u = \left(\sum_{i=1}^n w_i \cdot x_i \right) - \theta \quad (3.1)$$

$$y = g(u) \Rightarrow y = g \left[\left(\sum_{i=1}^n w_i \cdot x_i \right) - \theta \right] \quad (3.2)$$

3.1 Funções de Ativação

Segundo Silva, Spatti e Flauzino, (2010), pode-se fazer a divisão das funções de ativação em dois grupos principais: funções parcialmente diferenciáveis e funções totalmente diferenciáveis.

1. Funções de ativação parcialmente diferenciáveis são as que apresentam pontos no quais as derivadas de primeira ordem não existem. As três principais funções desse grupo são: função degrau, função degrau bipolar e função rampa simétrica.

- a) Função degrau (*heavyside / hard limiter*): o resultado gerado através da utilização da função degrau terá valores unitários positivos, assim que o potencial de ativação do neurônio for maior ou igual a zero, do contrário o resultado apresentará valores nulos, como apresentado na Equação 3.3 e ilustrado na Figura 3.3.

$$g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases} \quad (3.3)$$

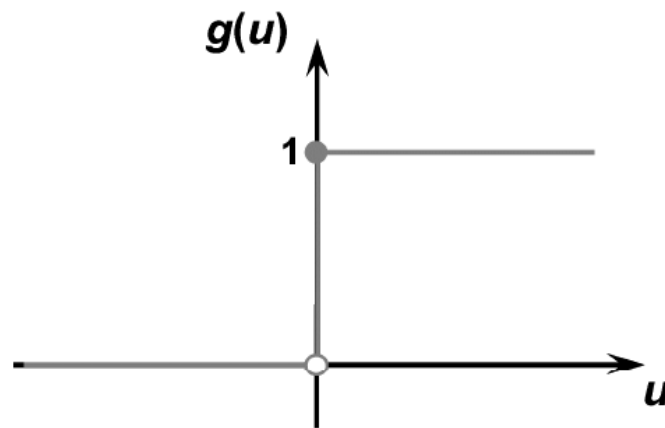


Figura 3.3: Função de ativação degrau.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

- b) Função degrau bipolar ou função sinal (*symmetric hard limiter*): o resultado fornecido pela utilização desta função irá apresentar valores positivos, no caso de o potencial de ativação do neurônio for maior que zero, valor nulo se o potencial também for nulo e valores negativos, se o potencial for menor do que zero. Essa função é apresentada na Figura 3.4 e na Equação 3.4.

$$g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u = 0 \\ -1, & \text{se } u \leq 0 \end{cases} \quad (3.4)$$

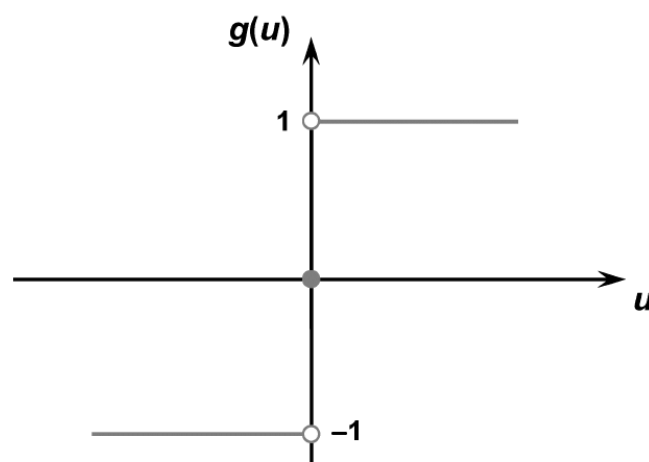


Figura 3.4: Função de ativação degrau bipolar.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

- c) Função rampa simétrica: os valores gerados são iguais aos valores apresentados pelos potenciais de ativação, se eles estão estabelecidos no intervalo $[-a, a]$, sendo limitados aos valores limites em caso contrário. A Figura 3.5 e a Equação 3.5 mostram essa função.

$$g(u) = \begin{cases} a, & \text{se } u > a \\ u, & \text{se } -a \leq u \leq a \\ -a, & \text{se } u < -a \end{cases} \quad (3.5)$$

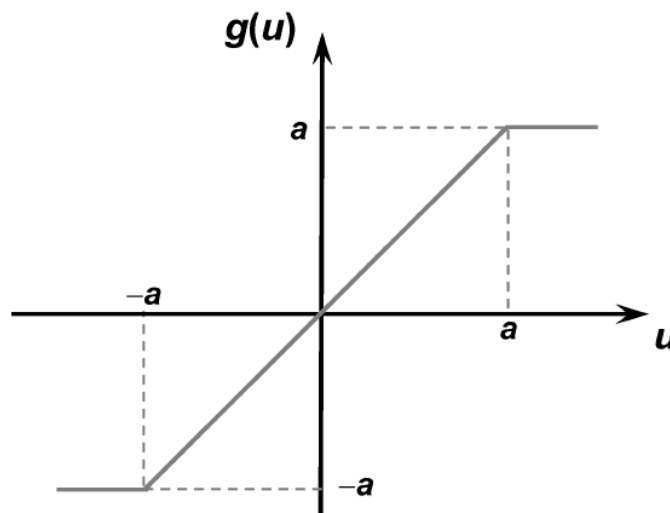


Figura 3.5: Função de ativação rampa simétrica.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

2. Funções de ativação totalmente diferenciáveis, são aquelas onde as derivadas de primeira ordem existem, e se conhece elas para cada ponto do seu domínio de definição. As quatro funções que fazem parte desse grupo e que são utilizadas em Redes Neurais Artificiais são a função logística, a tangente hiperbólica, a gaussiana e a função linear.

- a) Função logística: o resultado de saída que é dado por meio da utilização da função logística irá receber valores reais entre 0 e 1, como é mostrado na Equação 3.6 e na Figura 3.6. Na Equação 3.6,

o θ é uma constante real que determina o nível de inclinação da função logística em referência ao ponto de inflexão dela.

$$g(u) = \frac{1}{1 + e^{-\theta u}} \quad (3.6)$$

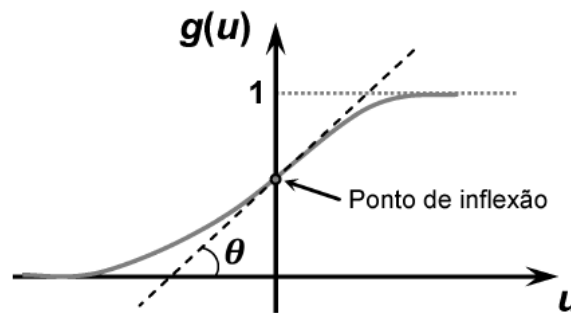


Figura 3.6: Função de ativação logística.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

- b) Função tangente hiperbólica: o resultado que é dado por esta função, receberá valores reais entre -1 e 1, e é apresentado na Equação 3.7 e ilustrado pela Figura 3.7. Nessa equação, tem-se que θ também está relacionado ao nível de inclinação da função tangente hiperbólica em referência ao ponto de inflexão dessa.

$$g(u) = \frac{1 - e^{-\theta u}}{1 + e^{-\theta u}} \quad (3.7)$$

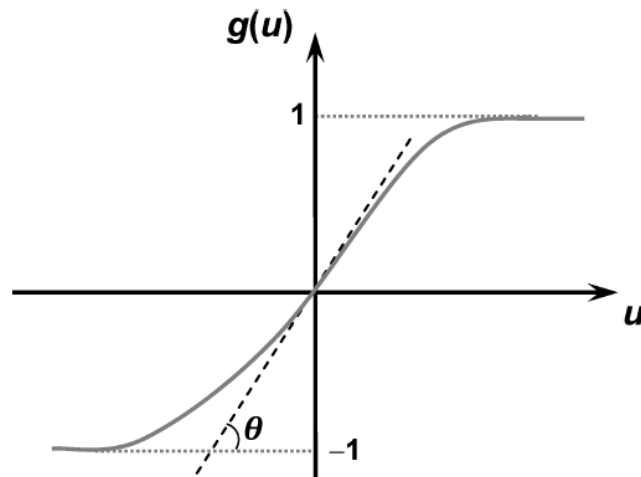


Figura 3.7: Função de ativação tangente hiperbólica.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

- c) Função gaussiana: o resultado obtido na função gaussiana, terá valores iguais, nos casos em que os valores de potencial de ativação estiverem localizados a uma mesma distância média de seu centro, já que a curva é simétrica ao centro. A Equação 3.8 mostra essa função, nela se tem o parâmetro c que representa o centro da função gaussiana, e σ o desvio padrão relacionado a curva. Na Figura 3.8, tem-se a ilustração dessa função.

$$g(u) = e^{-\frac{(u-c)^2}{2\sigma^2}} \quad (3.8)$$

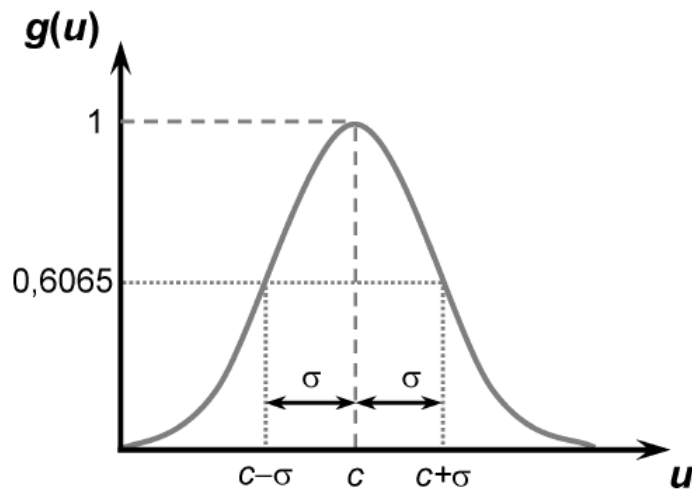


Figura 3.8: Função de ativação gaussiana.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

- d) Função linear ou função identidade: gera resultados de saídas iguais aos valores do potencial de ativação u . A Equação 3.9 demonstra ela matematicamente e a Figura 3.9 ilustra ela graficamente.

$$g(u) = u \quad (3.9)$$

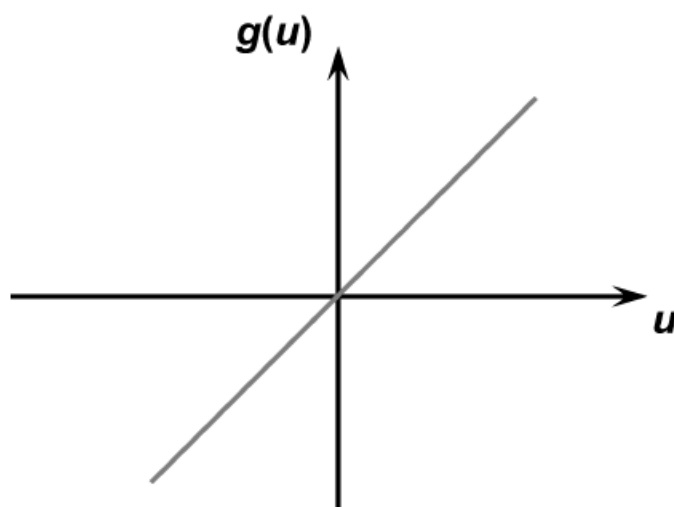


Figura 3.9: Função de ativação linear.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

3.2 Arquiteturas de Redes Neurais Artificiais

A arquitetura de uma RNA estabelece como os seus neurônios estão organizados uns em relação aos outros. Considerando-se a disposição dos neurônios, a forma que eles se interligam e como as camadas são constituídas, pode-se dividir as arquiteturas de RNA em: redes *feedforward* (alimentação à frente) de camada simples, redes *feedforward* de camadas múltiplas, redes recorrentes e redes reticuladas (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

3.2.1 Arquitetura *Feedforward* de Camadas Simples

Este tipo de arquitetura de RNAs contém somente uma camada de entrada e só uma camada de neurônios, que é a própria camada de saída. A Figura 3.10 apresenta uma rede *feedforward* de n camadas de entrada e m camadas de saída. As informações fluem em uma só direção, que é da camada de entrada para a camada de saída. Pode-se observar na Figura 3.10, que as quantidades de saídas são iguais aos neurônios de saída. Entre as redes mais utilizadas desse tipo de arquitetura, tem-se a *Perceptron* e *Andaline* (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

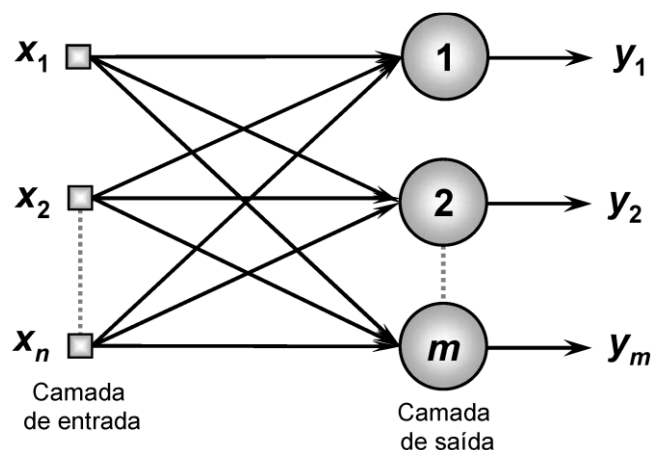


Figura 3.10: Arquitetura *feedforward* de camada simples

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

3.2.2 Arquitetura *Feedforward* de Múltiplas Camadas

As redes *feedforward* de múltiplas camadas são formadas por uma ou mais camadas escondidas de neurônios. Na Figura 3.11 apresenta-se uma rede *feedforward* de camadas múltiplas, que é composta por uma camada de entrada formada por n sinais, duas camadas escondidas formadas por n_1 e n_2 neurônios respectivamente e uma camada neural de saída constituída por m neurônios. O *Perceptron* multicamadas e a Rede Neural de base radial estão entre as principais redes desse tipo (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

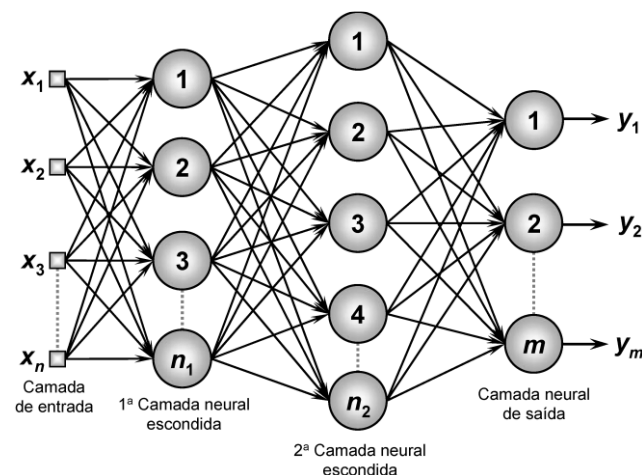


Figura 3.11: Arquitetura *feedforward* de camadas múltiplas.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

3.2.3 Arquitetura Recorrente ou Realimentada

São as redes nas quais as saídas dos neurônios são reaproveitadas como sinais de entrada por outros neurônios. A rede de *Hopfield* e a rede *Perceptron* multicamadas com realimentação entre os neurônios situados entre camadas distintas, são as mais utilizadas desse tipo de arquitetura. A Figura 3.12 apresenta um exemplo dessa arquitetura (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

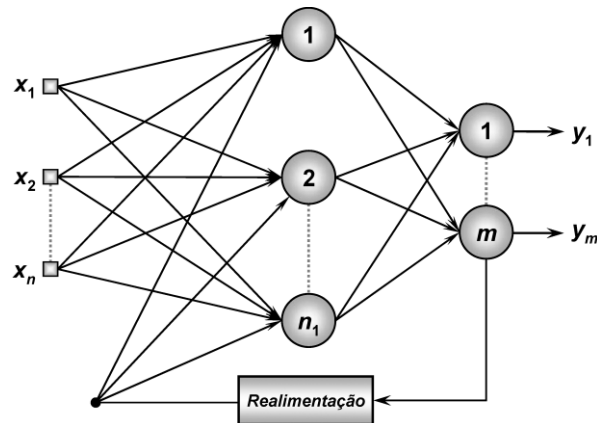


Figura 3.12: Arquitetura recorrente ou realimentada.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

3.2.4 Arquitetura em Estrutura Reticulada

Essa rede tem por características principais que os seus neurônios são dispostos espacialmente considerando-se o processo de ajuste de pesos e limiares. São utilizadas em problemas de agrupamento, reconhecimento de padrões, grafos, dentre outros. A rede de *Kohonen* é a mais utilizada desse tipo de arquitetura. Na Figura 3.13 é mostrado um exemplo de como é a arquitetura em estrutura reticulada (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

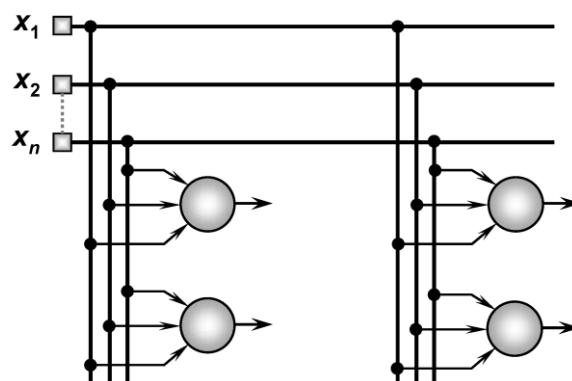


Figura 3.13: Arquitetura em estrutura reticulada.

Fonte: Silva, Spatti e Flauzino, 2010; Pereira, 2018.

3.3 Tipos de Treinamento para Rede Neural Artificial

Uma importante característica das Redes Neurais é que elas são capazes de aprender através de amostras (padrões), que demonstram o comportamento do sistema. Depois que a rede aprende a relação entre as entradas e saídas, ela é capaz de generalizar soluções. Sendo assim, através dos próximos sinais de entrada, a rede neural pode gerar uma saída próxima daquela esperada.

Para realizar o aprendizado de uma rede neural é realizado o treinamento desta. Para a realização deste treinamento é utilizado o algoritmo de aprendizagem, que são passos ordenados com a finalidade de generalizar as soluções de saída (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

Frequentemente, o conjunto de amostras à disposição sobre o comportamento do sistema divide-se em dois subconjuntos, o subconjunto de treinamento e o subconjunto de teste. No subconjunto de treinamento escolhe-se aleatoriamente de 60% a 90 % do total de amostras, estas que serão utilizadas para o processo de aprendizagem da rede. O subconjunto teste geralmente é composto por 10% a 40% das amostras totais, este é utilizado para analisar se a generalização de soluções da rede está em padrões aceitáveis (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

Enquanto o treinamento ocorre, cada vez que todas as amostras do subconjunto treinamento forem apresentadas e todas as etapas de treinamento ocorrerem, será chamada de época de treinamento. Existem vários tipos de treinamento de RNA, porém iremos descrever o treinamento supervisionado, utilizado neste trabalho.

O treinamento supervisionado é aquele no qual os sinais de entrada possuem os seus respectivos sinais de saída esperados. Por outro lado, o treinamento não supervisionado é aquele que não há as saídas desejadas (SILVA; SPATTI; FLAUZINO, 2010; PEREIRA, 2018).

3.4 Redes *Perceptron* Multicamadas

As redes *Perceptron* de Múltiplas Camadas ou *Perceptron* Multicamadas (PMC), tem como aspecto principal apresentarem ao menos uma camada intermediária ou escondida de neurônios, localizada no meio das camadas de entrada

e de saída neurais. Sendo assim, esse tipo de rede possui ao menos duas camadas de neurônios, uma intermediária e uma de saída. As redes PMC podem ser aplicadas em diversos tipos problemas, as principais áreas de aplicação são:

- Aproximação universal de funções;
- Reconhecimento de padrões;
- Identificação e controle de processos;
- Previsão de séries temporais;
- Otimização de sistemas.

A arquitetura das redes PMC são do tipo *feedforward* de camadas múltiplas. O fluxo das informações é feito começando pela camada de entrada, passa-se pelas camadas intermediárias e termina na camada neural de saída. A Figura 3.14 mostra o exemplo de uma rede PMC (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

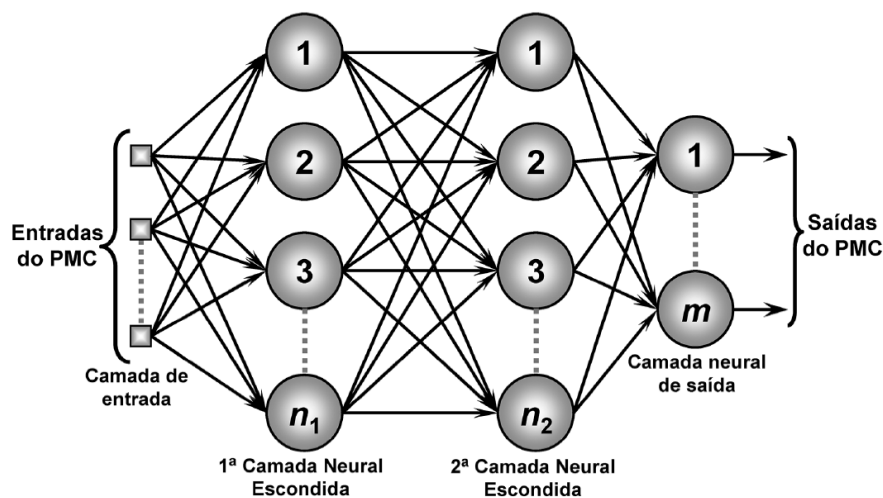


Figura 3.14: Exemplo de rede *Perceptron* multicamadas.

Fonte: Silva, Spatti e Flauzino, 2016.

3.4.1 Funcionamento da rede *Perceptron* multicamadas

Através da Figura 3.14, pode-se observar que cada uma das entradas da rede, que representam os sinais que vêm de uma determinada aplicação, irá ser propagada uma a uma até a camada neural de saída do PMC. Sendo assim, as saídas dos neurônios da 1ª camada neural, irão ser as entradas dos neurônios da segunda

camada neural escondida. Na representação da Figura 3.14, as entradas da camada neural de saída, serão as saídas da 2ª camada neural escondida.

Dessa forma, os sinais de entrada numa rede *Perceptron* multicamadas são propagados, independente de quantas camadas escondidas há, somente em um sentido, iniciando na camada de entrada e terminando na camada neural de saída (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

Para o funcionamento deste tipo de rede é feita a distribuição do conhecimento de como as entradas e saídas do sistema se comportam através de todos os neurônios que formam o PMC. Na camada de entrada são entregues os estímulos ou sinais para a rede. Por meio das camadas intermediária são extraídas a maioria das informações sobre os estímulos que são codificados através dos pesos sinápticos e limiares de seus neurônios, criando desta maneira, um modelo do ambiente no qual o sistema que está sendo tratado, está inserido. Por último, os estímulos provenientes dos neurônios da última camada intermediária são recebidos pelos neurônios da camada de saída, gerando um padrão de resposta que irá ser a saída que a rede vai disponibilizar (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

3.4.2 Processo de treinamento do *Perceptron* multicamadas

Para se realizar o treinamento de redes PMC, utiliza-se neste trabalho o algoritmo *backpropagation*, também chamado de regra Delta generalizada. São feitas aplicações sucessivas de duas fases bem específicas. A Figura 3.15 mostra essas duas fases para uma rede PMC que possui duas camadas escondidas. Essa rede possui n sinais na sua camada de entrada, n_1 neurônios na 1ª camada neural escondida, n_2 neurônios na 2ª camada neural escondida e n_3 sinais em sua camada de saída (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

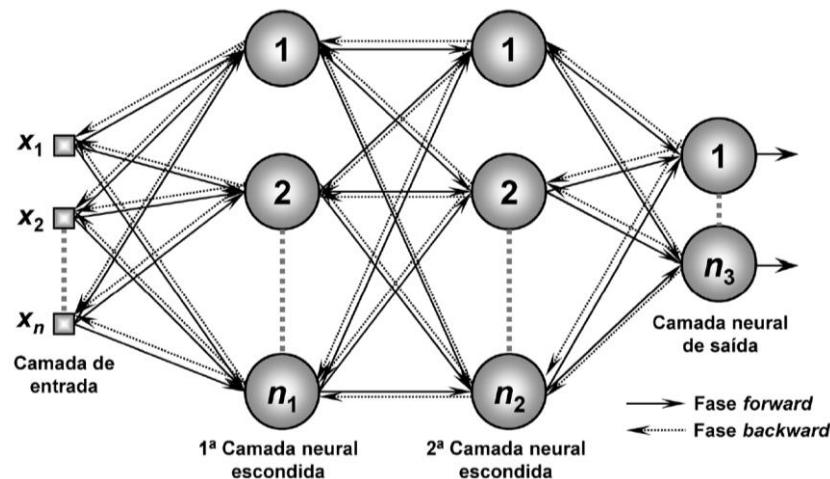


Figura 3.15: As duas fases de treinamento da rede PMC.

Fonte: Silva, Spatti e Flauzino, 2016.

A primeira fase de treinamento é chamada de “propagação adiante” (*forward*), nela é feita a inserção dos sinais ($x_1, x_2, \dots, x_n$) de uma amostra do conjunto de treinamento, nas entradas da rede é feita a propagação camada por camada até se obter as respectivas saídas. Essa fase tem por objetivo obter as respostas da rede, utilizando-se os valores atuais de pesos sinápticos e limiares de seus neurônios, pois estes não se alteram cada vez que essa fase é executada (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

Logo após se terem respostas nas saídas da rede, é feita uma comparação destas com as respectivas respostas esperadas, pois este é um processo de aprendizagem supervisionada. Para essa rede mencionada na Figura 3.15, que possui n_3 neurônios na camada de saída, é realizado o cálculo dos n_3 desvios (erros) entre cada uma das respostas esperadas e as que foram calculadas. Através dos erros resultantes é realizado o ajuste de pesos e limiares de todos os neurônios (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

Dessa forma, por meio desses valores de erros, realiza-se, logo após, a segunda fase do método *backpropagation*, chamada de “propagação reversa” (*backward*). De forma diferente a fase anterior, os pesos sinápticos e limiares são ajustados durante a execução dessa fase. Resumidamente, através de aplicações sucessivas das fases *forward* e *backforward*, os pesos sinápticos e limiares dos neurônios são ajustados de forma automática cada vez que uma iteração é realizada, fazendo com que a soma dos erros produzidos pelas respostas da rede em relação

às desejadas, diminua pouco a pouco. (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

3.4.3 O Algoritmo *Backpropagation*

A fim, de que se possa entender melhor o princípio de funcionamento do algoritmo *backpropagation*, é necessário primeiramente realizar a definição de várias variáveis e parâmetros auxiliares que serão utilizados. Por meio da topologia de PMC apresentada na Figura 3.15, a Figura 3.16 mostra um conjunto de variáveis que norteiam a derivação do algoritmo.

A configuração de cada neurônio (j) que pertence a uma das camadas (L) é mostrada na Figura 3.17, onde $g(\cdot)$ é uma função de ativação que tem que ser contínua e diferenciável em todo seu domínio. (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

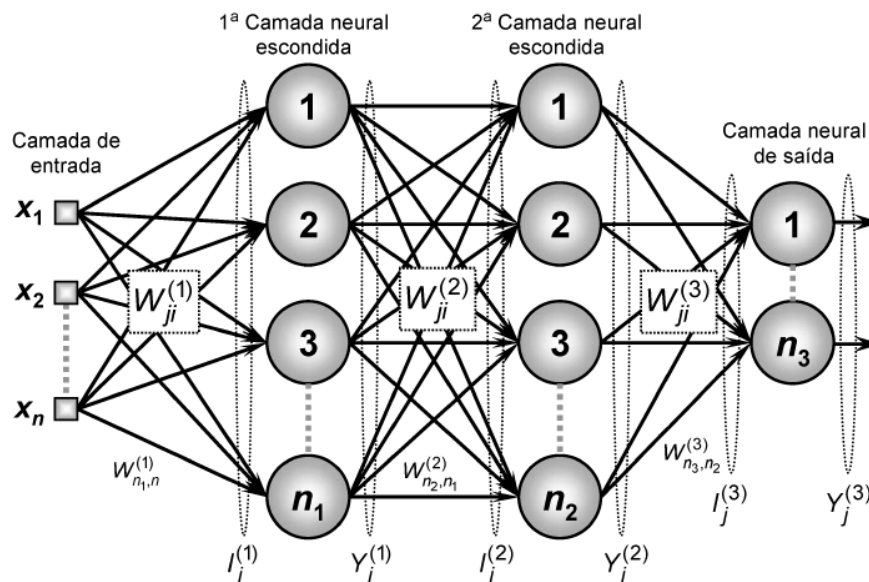


Figura 3.16: Notação para derivação do algoritmo *backpropagation*.

Fonte: Silva, Spatti e Flauzino, 2016.

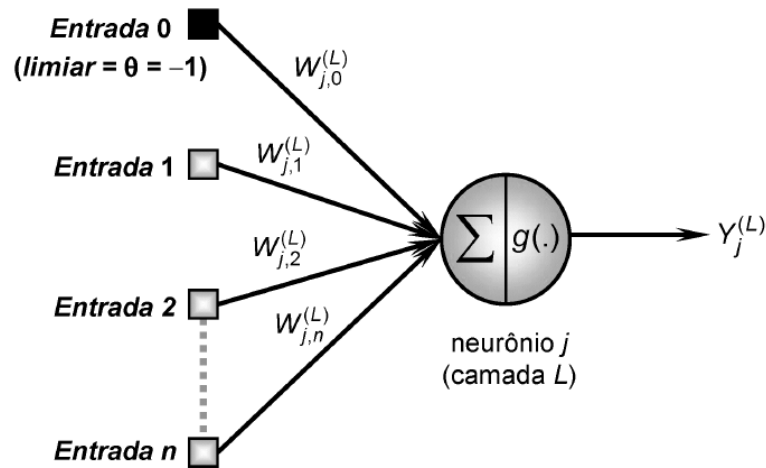


Figura 3.17: Configuração de neurônio usado na derivação do algoritmo *backpropagation*

Fonte: Silva, Spatti e Flauzino, 2016.

Através das figuras 3.16 e 3.17, tem-se a seguinte terminologia para os parâmetros do PMC:

- $W_{ji}^{(L)}$ são matrizes de pesos, onde os elementos representam o valor do peso sináptico ligando j -ésimo neurônio da camada (L) ao i -ésimo neurônio da camada ($L-1$). No caso mostrado na figura 3.16, tem-se:
 - $W_{ji}^{(3)}$ é o peso sináptico que liga o j -ésimo neurônio da camada de saída ao i -ésimo neurônio da camada 2;
 - $W_{ji}^{(2)}$ é o peso sináptico que conecta o j -ésimo neurônio da camada escondida 2 ao i -ésimo neurônio da camada 1;
 - $W_{ji}^{(1)}$ é o peso sináptico que faz a conexão entre o j -ésimo neurônio da camada 1 ao i -ésimo sinal da camada de entrada.
- $I_j^{(L)}$ são vetores no qual os elementos representam a entrada ponderada em relação ao j -ésimo neurônio da camada L , estes que são mostrados nas Equações (3.10), (3.11) e (3.12):

$$I_j^{(1)} = \sum_{i=0}^n W_{ji}^{(1)} \cdot x_i \Leftrightarrow I_j^{(1)} = W_{1,0}^{(1)} \cdot x_0 + W_{1,1}^{(1)} \cdot x_1 + \dots + W_{1,n}^{(1)} \cdot x_n \quad (3.10)$$

$$I_j^{(2)} = \sum_{i=0}^{n_1} W_{ji}^{(2)} \cdot Y_i^{(1)} \Leftrightarrow I_j^{(2)} = W_{1,0}^{(2)} \cdot Y_0^{(1)} + W_{1,1}^{(2)} \cdot Y_1^{(1)} + \dots + W_{1,n_1}^{(2)} \cdot Y_{n_1}^{(1)} \quad (3.11)$$

$$I_j^{(3)} = \sum_{i=0}^{n_2} W_{ji}^{(3)} \cdot Y_i^{(2)} \Leftrightarrow I_j^{(3)} = W_{1,0}^{(3)} \cdot Y_0^{(2)} + W_{1,1}^{(3)} \cdot Y_1^{(2)} + \dots + W_{1,n_2}^{(3)} \cdot Y_{n_2}^{(2)} \quad (3.12)$$

$Y_j^{(L)}$ são vetores em que seus elementos mostram a saída do j -ésimo neurônio, em relação à camada L e são dados nas Equações (3.13), (3.14), (3.15):

$$Y_j^{(1)} = g(I_j^{(1)}) \quad (3.13)$$

$$Y_j^{(2)} = g(I_j^{(2)}) \quad (3.14)$$

$$Y_j^{(3)} = g(I_j^{(3)}) \quad (3.15)$$

Agora, para começar a derivação do algoritmo *backpropagation* é feita a definição da função representativa do erro de aproximação, que irá mensurar o desvio entre as respostas obtidas pelos neurônios de saída da rede aos seus valores esperados (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018). Desse modo, levando em consideração a k -ésima amostra de treinamento na topologia da Figura 3.17, admite-se que a função de erro quadrático é a aquela que irá fazer medição do desempenho local ligado aos resultados obtidos pelos neurônios de saída comparados com a amostra relacionada, como mostrado na Equação (3.16):

$$E(k) = \frac{1}{2} \sum_{j=1}^{n_3} \left(d_j(k) - Y_j^{(3)}(k) \right)^2 \quad (3.16)$$

Dessa equação, tem-se que $Y_j^{(3)}(k)$ é o valor que o j -ésimo neurônio de saída da rede gera, levando-se em consideração a k -ésima amostra de treinamento, ao mesmo tempo que $d_j(k)$ é seu devido valor esperado.

Portanto, admitindo que o conjunto de treinamento possui p amostras, através da análise do “erro quadrático médio”, apresentado na Equação (3.17), pode-se mensurar o desempenho geral do algoritmo *backpropagation*, utilizando o $E(k)$ da Equação (3.16), que é o erro quadrático médio (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

$$E_M = \frac{1}{p} \sum_{k=1}^p E(k) \quad (3.17)$$

Para se compreender melhor os passos necessários para entender o algoritmo *backpropagation*, é feita a divisão de sua explicação em duas partes, na primeira é feito o ajuste dos pesos sinápticos da camada de saída e na segunda o ajuste de pesos sinápticos de camadas intermediárias. Para isto, utiliza-se de um método que se baseia no gradiente da função erro quadrático apresentada na Equação (3.16) (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

Parte I – Ajuste dos pesos sinápticos da camada de saída

Nesta parte é realizado o ajuste da matriz de pesos $W_{ji}^{(1)}$ para fazer a minimização do erro entre a saída gerada pela rede neural e sua saída desejada. Levando-se em consideração o erro que foi mostrado na Equação (3.16) que se relaciona com a k -ésima amostra de treinamento que se refere ao j -ésimo neurônio da camada de saída, e utilizando-se do conceito de gradiente e a regra da cadeia para derivadas parciais, obtêm-se a Equação (3.18):

$$\nabla E^{(3)} = \frac{\partial E}{\partial W_{ji}^{(3)}} = \frac{\partial E}{\partial Y_j^{(3)}} \cdot \frac{\partial Y_j^{(3)}}{\partial I_j^{(3)}} \cdot \frac{\partial I_j^{(3)}}{\partial W_{ji}^{(3)}} \quad (3.18)$$

A partir da Equação (3.12), obteve-se a Equação (3.19).

$$\frac{\partial I_j^{(3)}}{\partial W_{ji}^{(3)}} = Y_i^{(2)} \quad (3.19)$$

Por meio da Equação (3.15), tem-se a Equação (3.20).

$$\frac{\partial Y_j^{(3)}}{\partial I_j^{(3)}} = g'(I_j^{(3)}) \quad (3.20)$$

Através da Equação (3.16), obtém-se a Equação (3.21).

$$\frac{\partial E}{\partial Y_j^{(3)}} = -(d_j - Y_j^{(3)}) \quad (3.21)$$

Onde $g'(\cdot)$ é a derivada de primeira ordem da função de ativação utilizada. Fazendo a substituição das Equações (3.19), (3.20) e (3.21) na (3.18), se obtém a Equação (3.22):

$$\frac{\partial E}{\partial W_{ji}^{(3)}} = -(d_j - Y_j^{(3)}) \cdot g'(I_j^{(3)}) \cdot Y_i^{(2)} \quad (3.22)$$

Dessa forma, deve-se fazer o ajuste da matriz de pesos $W_{ji}^{(3)}$ em direção contrária ao gradiente para se minimizar o erro, conforme mostrado na Equação (3.23):

$$\Delta W_{ji}^{(3)} = -\eta \cdot \frac{\partial E}{\partial W_{ji}^{(3)}} \Leftrightarrow \Delta W_{ji}^{(3)} = \eta \cdot \delta_j^{(3)} \cdot Y_i^{(2)} \quad (3.23)$$

Da matriz de pesos, temos que $\delta_j^{(3)}$ é o gradiente que está relacionado com j-ésimo neurônio da camada de saída, ele é apresentado na Equação (3.24):

$$\delta_j^{(3)} = (d_j - Y_j^{(3)}) \cdot g'(I_j^{(3)}) \quad (3.24)$$

Pode-se converter a Equação (3.23), no procedimento iterativo da Equação (3.25):

$$W_{ji}^{(3)}(t+1) = W_{ji}^{(3)}(t) + \eta \cdot \delta_j^{(3)} \cdot Y_i^{(2)} \quad (3.25)$$

Da Equação (3.25) temos o η que representa a taxa de aprendizagem do algoritmo *backpropagation*. Passando essa Equação (3.25) para notação algorítmica, obtém-se a Equação (3.26):

$$W_{ji}^{(3)} \leftarrow W_{ji}^{(3)} + \eta \cdot \delta_j^{(3)} \cdot Y_i^{(2)} \quad (3.26)$$

Sendo assim, a Equação (3.26), realiza o ajuste dos pesos dos neurônios da camada de saída da rede, considerando-se a diferença que existe entre as respostas resultantes de suas saídas, e relacionando-as com seus respectivos valores esperados (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

Parte II – Ajuste dos pesos sinápticos de camadas intermediárias

Não se possui acesso direto aos valores esperados pelas saídas das camadas intermediárias, o que é bem diferente do que acontece com os neurônios das camadas de saída do PMC. Dessa forma, faz-se o ajuste de pesos sinápticos através das aproximações dos erros de saída que resultam dos neurônios da camada que vem logo em seguida, que já foram ajustados antecipadamente (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

Prosseguindo com a sequência de ajustes que são exigidos pela fase *backward* e usando a topologia da Figura 3.17. As correções dos pesos das camadas intermediárias começam a ser realizadas, após os neurônios da camada de saída terem sido ajustados. Com essa condição em específico, não há disponível os valores esperados para as saídas destes neurônios, dessa forma os ajustes são feitos utilizando os valores ponderados dos pesos sinápticos que já passaram pelo processo

de ajuste para os neurônios da camada de saída (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

(A) Ajuste dos pesos sinápticos da segunda camada

Essa parte do processo de treinamento é feito para se ajustar a matriz de pesos $W_{ji}^{(2)}$, com o objetivo de minimizar o erro entre a saída que foi gerada pela rede, levando-se em consideração a retro propagação do erro, proveniente dos cálculos feitos pelos neurônios da camada neural de saída por meio da Equação (3.27).

$$\nabla E^{(2)} = \frac{\partial E}{\partial W_{ji}^{(2)}} = \frac{\partial E}{\partial Y_j^{(2)}} \cdot \frac{\partial Y_j^{(2)}}{\partial I_j^{(2)}} \cdot \frac{\partial I_j^{(2)}}{\partial W_{ji}^{(2)}} \quad (3.27)$$

A partir da Equação (3.18), obteve-se a Equação (3.28).

$$\frac{\partial I_j^{(2)}}{\partial W_{ji}^{(2)}} = Y_i^{(1)} \quad (3.28)$$

Por meio Equação (3.21), obtém-se a Equação (3.29).

$$\frac{\partial Y_j^{(2)}}{\partial I_j^{(2)}} = g'(I_j^{(2)}) \quad (3.29)$$

$$\frac{\partial E}{\partial Y_j^{(2)}} = \sum_{k=1}^{n_3} \frac{\partial E}{\partial I_k^{(3)}} \cdot \frac{\partial I_k^{(3)}}{\partial Y_j^{(2)}} = \underbrace{\sum_{k=1}^{n_3} \frac{\partial E}{\partial I_k^{(3)}}}_{\text{parte (i)}} \cdot \underbrace{\frac{\partial (\sum_{k=1}^{n_3} W_{kj}^{(3)} \cdot Y_j^{(2)})}{\partial Y_j^{(2)}}}_{\text{parte (ii)}} \quad (3.30)$$

Tem-se da Equação (3.30) que o valor de $W_{kj}^{(2)}$ é exatamente o valor da derivação parcial do argumento da parte (ii) em relação à $Y_j^{(2)}$, resultando na Equação (3.31).

$$\frac{\partial E}{\partial Y_j^{(2)}} = \sum_{k=1}^{n_3} \underbrace{\frac{\partial E}{\partial I_k^{(3)}}}_{\text{parte (i)}} \cdot \underbrace{W_{kj}^{(3)}}_{\text{parte (ii)}} \quad (3.31)$$

Faz-se necessário ressaltar que o valor da parte (ii) da Equação (3.31), é referente aos pesos sinápticos de todos os neurônios da camada de saída, que são conectados a um indicado neurônio j da segunda camada intermediária. Na etapa anterior realizou-se os ajustes de todos os pesos $W_{ji}^{(3)}$, baseando-se em valores reais de erro, esses que irão ser usados para ajustar os pesos da segunda camada neural intermediária (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

Para se obter os valores para a parte (i) da Equação (3.31), pode-se calcular o resultado dessa por meio da multiplicação de (3.20) por (3.21), resultando no respectivo valor de $\delta_j^{(3)}$ mostrado na Equação (3.24). Dessa forma, considerando-se estas substituições, a Equação (3.31) torna-se a Equação (3.32).

$$\frac{\partial E}{\partial Y_j^{(2)}} = - \sum_{k=1}^{n_3} \delta_k^{(3)} \cdot W_{kj}^{(3)} \quad (3.32)$$

Realizando-se a substituição das Equações (3.28), (3.29) e (3.32) na Equação (3.27), obtém-se a Equação (3.33).

$$\frac{\partial E}{\partial W_{ji}^{(2)}} = - \left(\sum_{k=1}^{n_3} \delta_k^{(3)} \cdot W_{kj}^{(3)} \right) \cdot g'(I_j^{(2)}) \cdot Y_i^{(1)} \quad (3.33)$$

Sendo assim, com a finalidade de se minimizar o erro, faz-se o ajuste da matriz de pesos $W_{ji}^{(2)}$ em direção contrária ao gradiente, ou seja:

$$\Delta W_{ji}^{(2)} = -\eta \cdot \frac{\partial E}{\partial W_{ji}^{(2)}} \Leftrightarrow \Delta W_{ji}^{(2)} = \eta \cdot \delta_j^{(2)} \cdot Y_i^{(1)} \quad (3.34)$$

onde $\delta_j^{(2)}$ é por definição o gradiente local em relação ao j -ésimo neurônio que pertence a segunda camada intermediária, que é mostrado na Equação (3.35).

$$\delta_j^{(2)} = \left(\sum_{k=1}^{n_3} \delta_k^{(3)} \cdot W_{kj}^{(3)} \right) \cdot g'(I_j^{(2)}) \quad (3.35)$$

A Equação (3.34) pode ser transformada no seguinte procedimento iterativo apresentado na Equação (3.36).

$$W_{ji}^{(2)}(t+1) = W_{ji}^{(2)}(t) + \eta \cdot \delta_j^{(2)} \cdot Y_i^{(1)} \quad (3.36)$$

Se a Equação (3.36) for expressa em notação algorítmica, tem-se a Equação (3.37).

$$W_{ji}^{(2)} \leftarrow W_{ji}^{(2)} + \eta \cdot \delta_j^{(2)} \cdot Y_i^{(1)} \quad (3.37)$$

Desse modo, a Equação (3.37) realiza o ajuste dos pesos dos neurônios pertencentes a segunda camada escondida da rede, considerando-se a retro propagação do erro proveniente dos neurônios da camada de saída.

(B) Ajuste dos pesos sinápticos da primeira camada escondida:

Para a primeira camada escondida, o processo de treinamento é realizado para fazer o ajuste da matriz de pesos $W_{ji}^{(2)}$, com o objetivo de se ter uma minimização do erro entre a saída que a rede produz, por meio da retro propagação do erro proveniente dos ajustes dos neurônios da segunda camada escondida. Dessa forma, tem-se a Equação (3.38):

$$\nabla E^{(1)} = \frac{\partial E}{\partial W_{ji}^{(1)}} = \frac{\partial E}{\partial Y_j^{(1)}} \cdot \frac{\partial Y_j^{(1)}}{\partial I_j^{(1)}} \cdot \frac{\partial I_j^{(1)}}{\partial W_{ji}^{(1)}} \quad (3.38)$$

A partir da Equação (3.10), obteve-se a Equação (3.39).

$$\frac{\partial I_j^{(1)}}{\partial W_{ji}^{(1)}} = x_i \quad (3.39)$$

Através da Equação (3.13), tem-se a Equação (3.40).

$$\frac{\partial Y_j^{(1)}}{\partial I_j^{(1)}} = g'(I_j^{(1)}) \quad (3.40)$$

$$\frac{\partial E}{\partial Y_j^{(1)}} = \sum_{k=1}^{n_2} \frac{\partial E}{\partial I_k^{(2)}} \cdot \frac{\partial I_k^{(2)}}{\partial Y_j^{(1)}} = \underbrace{\sum_{k=1}^{n_2} \frac{\partial E}{\partial I_k^{(2)}}}_{\text{parte (i)}} \cdot \underbrace{\frac{\partial (\sum_{k=1}^{n_2} W_{kj}^{(2)} \cdot Y_j^{(1)})}{\partial Y_j^{(1)}}}_{\text{parte (ii)}} \quad (3.41)$$

Na Equação (3.41) acontece semelhante à Equação (3.30), tem-se que o valor da derivação parcial do argumento da parte (ii) da Equação em relação $Y_j^{(1)}$ é representado pelo valor de $W_{kj}^{(2)}$, o que resulta na Equação (3.42).

$$\frac{\partial E}{\partial Y_j^{(1)}} = \sum_{k=1}^{n_2} \underbrace{\frac{\partial E}{\partial I_k^{(2)}}}_{\text{parte (i)}} \cdot \underbrace{W_{kj}^{(2)}}_{\text{parte(ii)}} \quad (3.42)$$

É importante destacar que o valor da parte (ii) da Equação (3.42) refere-se aos pesos sinápticos de todos os neurônios da segunda camada intermediária, que são conectados a um certo neurônio j da primeira camada intermediária. Na etapa anterior foram feitos os ajustes de todos os pesos $W_{ji}^{(2)}$, usando-se os valores retro propagados de erro, que são calculados no ajuste dos pesos $W_{ji}^{(3)}$, esses que foram ajustados baseando-se em valores reais de erro (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

Para se calcular os valores para a parte (i) da Equação (3.42) é necessário realizar a multiplicação de (3.29) por (3.30), resultando no respectivo valor de $\delta_j^{(2)}$

apresentado na Equação (3.35). Sendo assim, realizando-se estas substituições, a Equação (3.41) se transforma na Equação (3.43):

$$\frac{\partial E}{\partial Y_j^{(1)}} = - \sum_{k=1}^{n_2} \delta_k^{(2)} \cdot W_{kj}^{(2)} \quad (3.43)$$

Realizando-se a substituição das Equações (3.39), (3.40) e (3.43) em (3.38), tem-se a Equação (3.44).

$$\frac{\partial E}{\partial W_{ji}^{(1)}} = - \left(\sum_{k=1}^{n_2} \delta_k^{(2)} \cdot W_{kj}^{(2)} \right) \cdot g'(I_j^{(1)}) \cdot x_i \quad (3.44)$$

Sendo assim, deve-se fazer o ajuste da matriz de pesos em direção contrária ao gradiente para se minimizar o erro, como mostrado na Equação (3.45).

$$\Delta W_{ji}^{(1)} = -\eta \cdot \frac{\partial E}{\partial W_{ji}^{(1)}} \Leftrightarrow \Delta W_{ji}^{(1)} = \eta \cdot \delta_j^{(1)} \cdot x_i \quad (3.45)$$

onde $\delta_j^{(1)}$ é por definição o gradiente local em relação ao j-ésimo neurônio que pertence a primeira camada intermediária, sendo este estabelecido pela Equação (3.46):

$$\delta_j^{(1)} = \left(\sum_{k=1}^{n_2} \delta_k^{(2)} \cdot W_{kj}^{(2)} \right) \cdot g'(I_j^{(1)}) \quad (3.46)$$

Pode se converter a Equação (3.45), no procedimento iterativo da Equação 3.47:

$$W_{ji}^{(1)}(t+1) = W_{ji}^{(1)}(t) + \eta \cdot \delta_j^{(1)} \cdot x_i \quad (3.47)$$

Colocando-se a Equação (3.47) em notação algorítmica, tem-se então a Equação (3.48):

$$W_{ji}^{(1)} \leftarrow W_{ji}^{(1)} + \eta \cdot \delta_j^{(1)} \cdot x_i \quad (3.48)$$

Utiliza-se como critério de parada o cálculo do erro quadrático médio realizado por meio da Equação (3.17), utilizando-se todas as amostras que estão disponíveis para treinamento. A convergência do algoritmo ocorre logo que o erro quadrático médio entre duas épocas consecutivas for suficientemente pequeno, ou seja, como apresentado na Equação 3.49, onde o ε é a precisão necessária para que aconteça a convergência, precisão esta que depende de qual será a aplicação estudada pela rede (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

$$|E_M^{atual} - E_M^{anterior}| \leq \varepsilon \quad (3.49)$$

Nos casos em que a precisão proposta se torne inalcançável, pode-se usar a variável época como critério de parada para o PMC. Para isso, só é necessário que se limite o processo para quando o número específico de épocas seja atingido, valor esse que deve ter sido pré-estabelecido.

É importante destacar que as matrizes de pesos são ajustadas só na fase de treinamento da rede, na qual são aplicados os passos *forward* e *backfoward* com a finalidade de se fazer eventuais correções sinápticas necessárias. Enquanto isso, na fase de teste não é feito nenhum tipo de ajuste nos parâmetros internos da rede, sendo que nesse caso, só se faz a fase *forward* com a finalidade de gerar as saídas da rede (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

No problema estudado nesta dissertação, utilizou-se o algoritmo *backpropagation* com taxa de *momentum*, que é uma versão aperfeiçoada do método *backpropagation*. Para isso é inserido o termo de *momentum*, com o objetivo de ponderar quantitativamente a alteração das matrizes sinápticas que ocorreu entre duas iterações anteriores e sucessivas. Se levar em consideração os neurônios que pertencem à L -ésima camada, temos o algoritmo *backpropagation* normal, que é representado pela Equação (3.50), mas se o termo de *momentum* α for inserido, obtém-se a Equação (3.51) (SILVA; SPATTI; FLAUZINO, 2016; PEREIRA, 2018).

$$W_{ji}^{(L)}(t+1) = W_{ji}^{(L)}(t) + \eta \cdot \delta_j^{(L)} \cdot Y_i^{(L-1)} \quad (3.50)$$

$$W_{ji}^{(L)}(t+1) = W_{ji}^{(L)}(t) + \underbrace{\alpha \cdot (W_{ji}^{(L)}(t) - W_{ji}^{(L)}(t-1))}_{\text{termo de momentum}} + \underbrace{\eta \cdot \delta_j^{(L)} \cdot Y_i^{(L-1)}}_{\text{termo de aprendizagem}} \quad (3.51)$$

A seguir, estão os algoritmos a serem seguidos para se aplicar as fases de Treinamento e de Teste do algoritmo *backpropagation*.

Algoritmo 1: Fase de Treinamento do PMC

- 1: Obter o conjunto de amostras de treinamento $x^{(k)}$;
 - 2: Associar o vetor de saída desejada $d^{(k)}$ para cada amostra;
 - 3: Iniciar $W_{ji}^{(1)}$ e $W_{ji}^{(2)}$ e $W_{ji}^{(3)}$ com valores aleatórios pequenos;
 - 4: Especificar a taxa de aprendizagem η e precisão requerida ε ;
 - 5: Iniciar o contador de número de épocas ($\text{época} \leftarrow 0$);
 - 6: Repetir as instruções:
 - 7: $E_M^{\text{anterior}} \leftarrow E_M$;
 - 8: Para todas as amostras de treinamento ($x^{(k)}, d^{(k)}$), fazer:
 - 9: Obter $I_j^{(1)}$ e $Y_j^{(1)}$;
 - 10: Obter $I_j^{(2)}$ e $Y_j^{(2)}$;
 - 11: Obter $I_j^{(3)}$ e $Y_j^{(3)}$;
 - 12: Determinar $\delta_i^{(3)}$;
 - 13: Ajustar $W_{ji}^{(3)}$;
 - 14: Determinar $\delta_i^{(2)}$;
 - 15: Ajustar $W_{ji}^{(2)}$;
 - 16: Determinar $\delta_i^{(1)}$;
 - 17: Ajustar $W_{ji}^{(1)}$;
 - 18: Fim para
 - 19: Obter $Y_j^{(3)}$ ajustado;
 - 20: $E_M^{\text{atual}} \leftarrow E_M$;
 - 21: $\text{época} \leftarrow \text{época} + 1$;
 - 22: Até que $|E_M^{\text{atual}} - E_M^{\text{anterior}}| \leq \varepsilon$
 - 23: Fim
-

A seguir tem-se o algoritmo para se realizar a fase de teste do PMC.

Algoritmo 2: Fase de Teste do PMC

- 1: Obter uma amostra x ;
 - 2: Assumir $W_{ji}^{(1)}$, $W_{ji}^{(2)}$, $W_{ji}^{(3)}$ já ajustadas no treinamento;
 - 3: Execute as seguintes instruções:
 - 4: Obter $I_j^{(1)}$ e $Y_j^{(1)}$;
 - 5: Obter $I_j^{(2)}$ e $Y_j^{(2)}$;
 - 6: Obter $I_j^{(3)}$ e $Y_j^{(3)}$;
 - 7: Disponibilizar as saídas da rede, que são dadas pelos elementos de $Y_j^{(3)}$
 - 8: Fim
-

Capítulo 4

4 Resultados e discussão

Nesta pesquisa foram utilizados 121 exames para criação da rede neural e para testar o funcionamento dessa. Esses que foram cedidos pelo hospital de Santo Antônio de Pádua, e que foram avaliados por duas enfermeiras e um médico. Eles avaliaram se os exames apresentavam infecções bacterianas ou não. Após comparar cada uma das avaliações feitas pelos profissionais da saúde, pode-se perceber divergências entre os profissionais.

Com os dados dos exames foi possível construir uma RNA. Para a criação dela foram utilizados 17 parâmetros de entrada, que são: idade, sexo, hemácias, hemoglobina, hematócrito, VCM, HCM, CHMC, leucócitos, basófilos, eosinófilos, mielócitos, jovens, bastões, segmentados, linfócitos e monócitos que são os parâmetros do hemograma. Sendo o sexo classificado entre 0 e 1, utilizando-se o 0 para representar o sexo feminino e 1 para o sexo masculino.

Para criar a RNA foram utilizados 90 exames para a realização do treinamento da rede, que foram escolhidos de forma aleatória dentre os quais os 3 avaliadores apresentaram a mesma análise, cerca de 74 % dos exames totais, total este que corresponde a 121 exames. Em caso de se considerar apenas o total de exames sem divergência de avaliação, esse total de exames selecionados representam cerca de 80% dos exames sem divergência. Para o primeiro teste utilizou-se 23 exames, também escolhidos aleatoriamente representando 19 % dos exames totais, ou seja, foram analisados somente os exames sem divergência, que são aproximadamente 20 % desses. Para o segundo teste usou-se os 8 exames que foram avaliados de forma divergente, aproximadamente 7% do total de exames. Na Figura 4.1 é apresentado um desenho esquemático de como são as entradas e saídas das redes criadas.

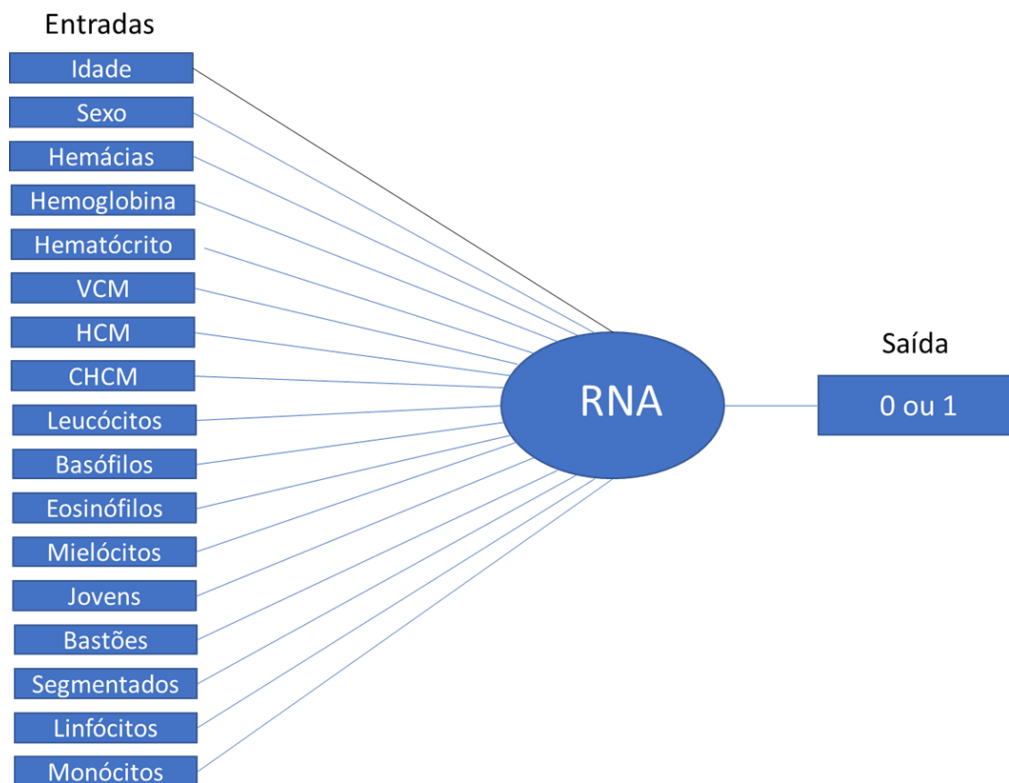


Figura 4.1: Desenho esquemático da RNA usada

Fonte: Autoria própria, 2023.

Para se treinar a rede utilizou-se o algoritmo *backpropagation* com momentum, implementado no *software Matlab*. Para a utilização do *Matlab*, usou-se um *Notebook Dell G7 7588* com processador Intel® Core™ i7-8750H, CPU @ 2.20 GHz a 4.10 GHz, 16 GB de memória *Ram* e Placa de Vídeo NVIDIA GeForce GTX 1060 with Max-Q Design com 6GB de Memória de Vídeo.

Na primeira camada que é a de entrada, foram utilizados 17 neurônios que é igual a quantidade de parâmetros de entrada, e a função de ativação escolhida para a primeira camada foi a *tansig*. Para o funcionamento da rede foram utilizadas duas camadas escondidas, que possuem a mesma quantidade de neurônios, para mostrar os diferentes desempenhos de treinamento mudou-se a quantidade de neurônios delas. Foram criadas 3 RNAs, a primeira com 100 neurônios nas camadas escondidas, a segunda com 200 e finalmente a terceira com cerca de 500 neurônios, e escolheu-se também a função de ativação *tansig*. Na camada de saída foi utilizado somente 1 neurônio, que fornecia as respostas através de 0, quando o exame não

possuía infecção bacteriana, e 1 apresentando que o exame apresentava infecção bacteriana, nessa camada utilizou-se a função *logsig*.

Os parâmetros da RNA foram: intervalo de interações mostradas igual a 50, taxa de aprendizagem de 0,0001, número máximo de épocas de treinamento igual a 20 mil, erro estipulado de 0,00001 e taxa de momentum de 0,7. Esses valores foram selecionados após a realização de vários testes, e observando o comportamento dos treinamentos das redes treinadas.

Antes de apresentar os resultados obtidos pelos treinamentos das RNAs, para se ter uma melhor ideia e interpretar adequadamente o desempenho destas, foi feito um gráfico, a fim de se entender melhor essa amostra utilizada para treinamento e testes das Redes Neurais criadas nessa pesquisa.

Considerando a amostra total, pode-se perceber pelo gráfico apresentado na Figura 4.2, que a maioria dos dados apresentados são de pacientes de sexo feminino que correspondem a cerca de 62 % dessa amostra, enquanto que os pacientes do sexo masculino são aproximadamente 38 %. Percebendo assim, que em relação a distribuição de sexo, não se tem uma discrepância tão grande nesse total.

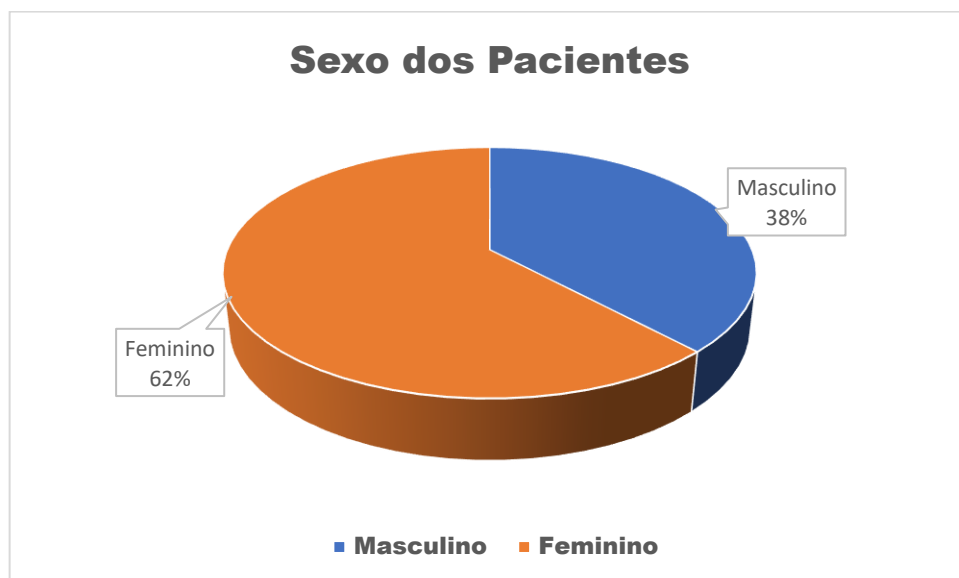


Figura 4.2: Gráfico de distribuição do sexo dos pacientes

Fonte: Autoria própria, 2023.

Foram feitos vários treinamentos da RNA utilizando 100 neurônios nas camadas intermediárias, até que se chegou no treinamento que foi apresentado na Figura 4.3. Devido as características computacionais e pequeno números de épocas,

além da quantidade de neurônios utilizados, esse treino foi bem rápido, sendo realizado em apenas 1 segundo. O treinamento final terminou quando se atingiu o erro de 0,00001.

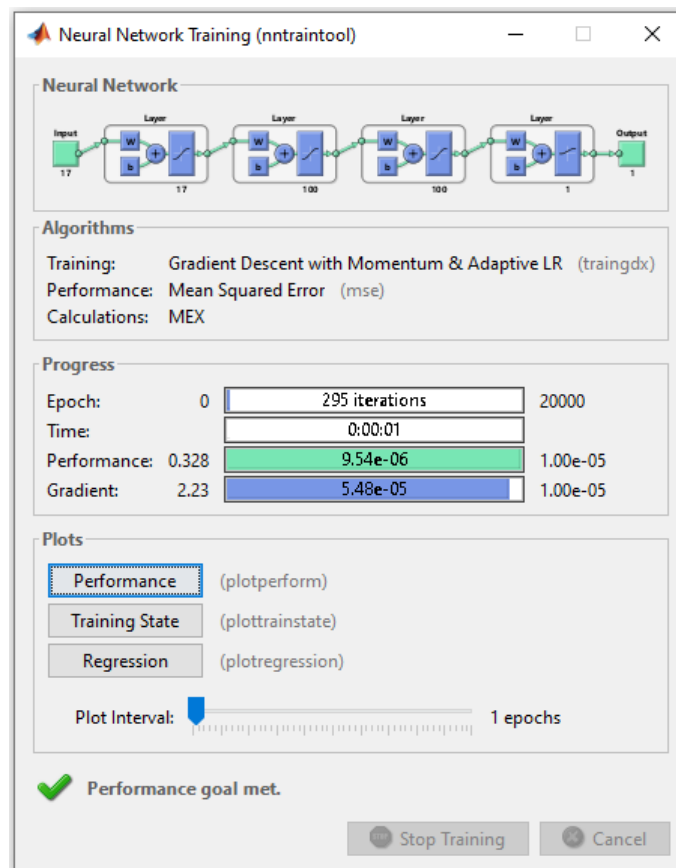


Figura 4.3: Treinamento da RNA com 100 neurônios.

Fonte: Autoria própria, 2023.

Na construção do código foi adicionado, a fim de se observar o desempenho do mesmo, a construção de um gráfico de barras, que é apresentado na Figura 4.4. Neste, as colunas da linha 1 são os resultados esperados pelos testes de acordo com as opiniões dos profissionais da saúde, no qual esperava-se que a RNA tivesse um comportamento o mais próximo possível desses profissionais. Já as colunas da linha 2 são análises feitas pela RNA.

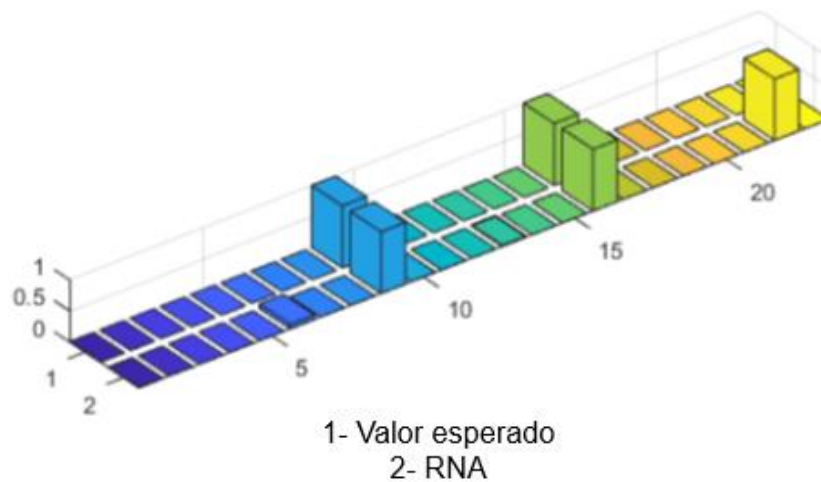


Figura 4.4: Gráfico de valor esperado x RNA com 100 neurônios

Fonte: Autoria própria, 2023.

Analisando o gráfico gerado, a RNA só teve uma interpretação diferente que foi o exame de número 22, que segundo o que havia sido avaliado, era para ser considerado como negativo, mas ao contrário do que era esperado, a rede interpretou como sendo positivo. Para uma melhor análise da Figura 4.4, criou-se também a Tabela 3, para poder observar os dados numericamente.

Tabela 3. Valor Esperado x RNA com 100 neurônios

Exame	Esperado	Realizado	Exame	Esperado	Realizado
1	0	0,005800	13	0	0,019200
2	0	0,000088	14	0	0,000378
3	0	0,000002	15	0	0,003300
4	0	0,000014	16	1	0,997300
5	0	0,000010	17	0	0,000016
6	0	0,076100	18	0	0,000105
7	0	0,000009	19	0	0,000056
8	0	0,000024	20	0	0,000078
9	1	0,993800	21	0	0,002300
10	0	0,005800	22	0	0,991100
11	0	0,001100	23	0	0,001400
12	0	0,000014			

Fonte: Autoria própria, 2023.

Tendo sido realizada a criação da segunda RNA, que possui 200 neurônios nas camadas intermediárias, o treinamento desta pode ser observado na Figura 4.5 e foi feito bem rapidamente com apenas 2 segundos e com um número de épocas dessa vez um pouco maior, totalizando 358 iterações, alcançando o erro estipulado novamente.

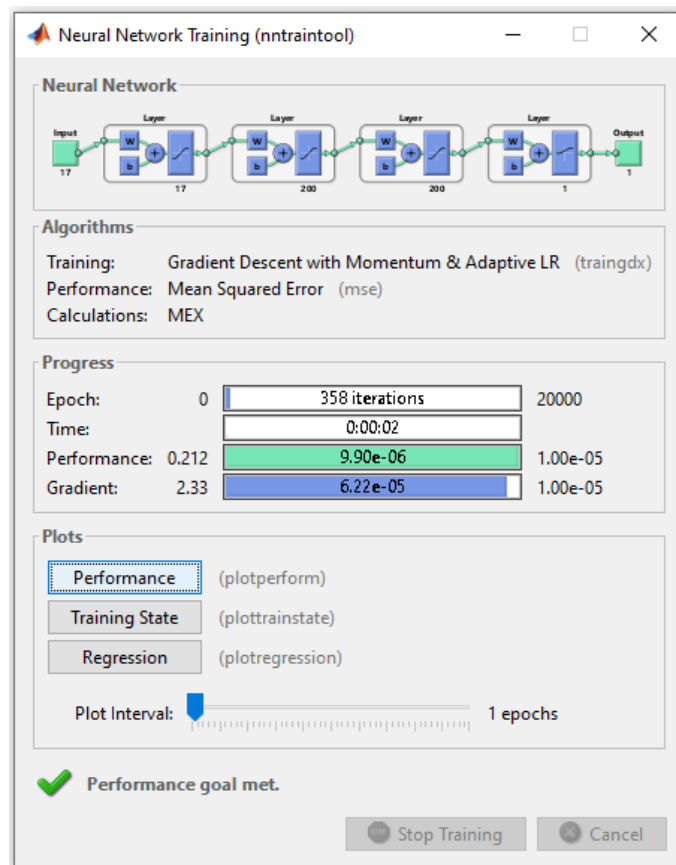


Figura 4.5: Treinamento da RNA com 200 neurônios

Fonte: Autoria própria, 2023.

Através do gráfico mostrado na Figura 4.6, pode-se inferir que essa RNA teve um excelente desempenho, se comportando totalmente igual as avaliações feitas pelos profissionais, tendo uma pequena diferença apenas na análise do exame 22.

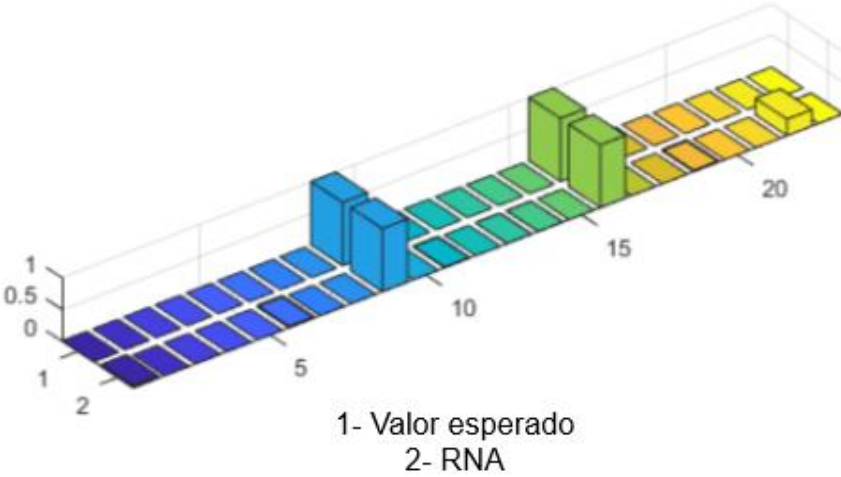


Figura 4.6: Gráfico de valor esperado x RNA com 200 neurônios

Fonte: Autoria própria, 2023.

A partir da Tabela 4, pode-se perceber a pequena diferença no exame 22, que apresentou 0,260700, que é próximo de 0, indicando assim que esse exame não apresenta infecção bacteriana, auxiliando na análise da Figura 4.6.

Tabela 4. Valor Esperado x RNA com 200 neurônios

Exame	Esperado	Realizado	Exame	Esperado	Realizado
1	0	0,043300	13	0	0,006600
2	0	0,000021	14	0	0,006400
3	0	0,000015	15	0	0,001700
4	0	0,000001	16	1	0,991300
5	0	0,000004	17	0	0,000105
6	0	0,028900	18	0	0,000169
7	0	0,000019	19	0	0,024000
8	0	0,000254	20	0	0,000569
9	1	0,999000	21	0	0,000204
10	0	0,000123	22	0	0,260700
11	0	0,015600	23	0	0,004000
12	0	0,000014			

Fonte: Autoria própria, 2023.

Finalmente a última RNA treinada, esta que possui 500 neurônios. como mostrado na Figura 4.7.

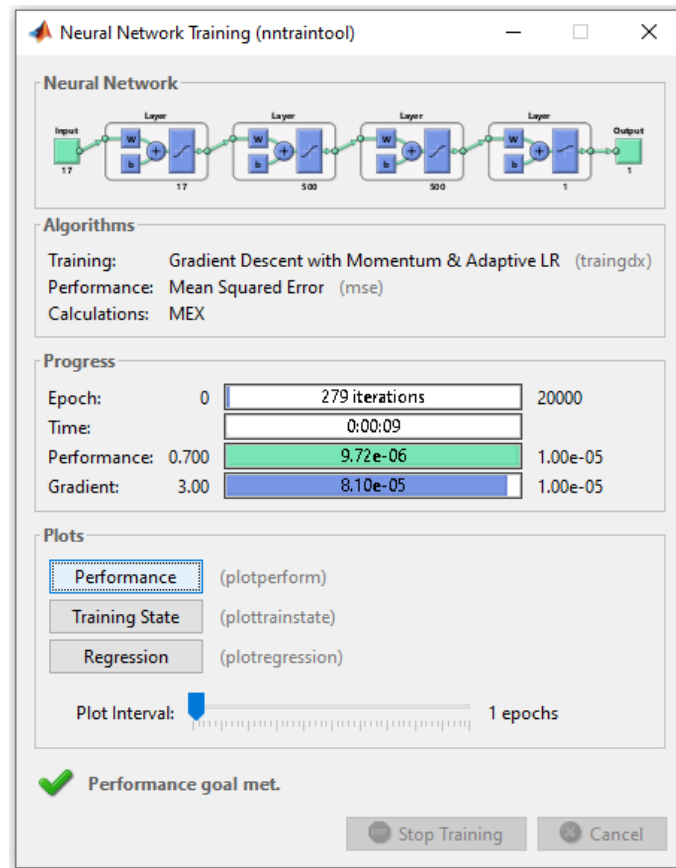


Figura 4.7: Treinamento da RNA com 500 neurônios

Fonte: Autoria própria, 2023.

Essa RNA precisou de menos iterações que a RNA com 200 neurônios para atingir o resultado esperado, demorando um pouco mais de tempo, por volta de 9 segundos para realizar o treino completo, mesmo assim, o tempo ainda é bem rápido, pois as configurações computacionais ajudaram a ter mais velocidade de treinamento. O aumento no tempo de treinamento ocorre pois foram utilizados mais neurônios. Dependendo da capacidade computacional poderia demorar muito mais de 1 ou 2 minutos para realizar esse treinamento.

Fazendo a análise do gráfico da Figura 4.8, pode-se perceber um ótimo desempenho no comportamento da RNA com 500 neurônios, acertando todas as avaliações e se comportando exatamente igual as avaliações dos 3 profissionais. Diferentemente das RNAs anteriores, esta teve pequenas diferenças avaliando os exames 10 e 13, e não teve se analisar o exame 22.

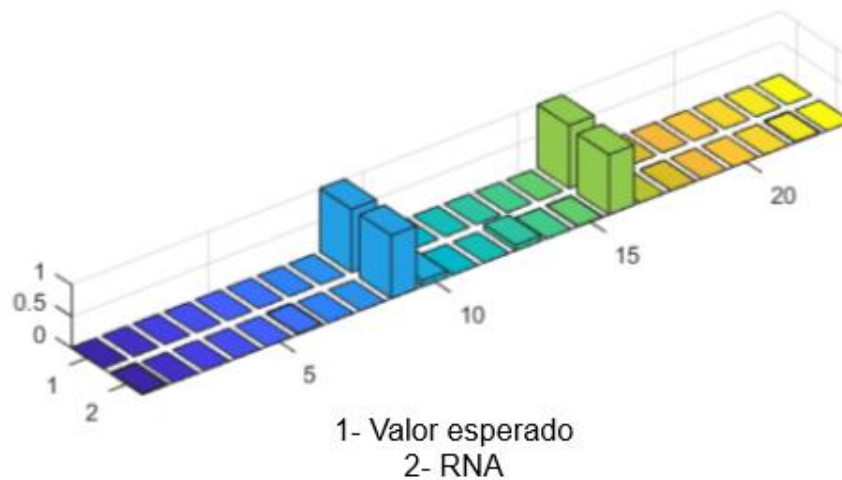


Figura 4.8: Gráfico de valor esperado x RNA para 500 neurônios

Fonte: Autoria própria, 2023.

Analisando a Figura 4.8 juntamente com os dados fornecidos pela Tabela 5, pode-se reforçar o que já foi citado, que das 3 RNAs criadas essa teve um comportamento mais próximo da análise dos 3 profissionais da medicina.

Tabela 5. Valor Esperado x RNA com 500 neurônios

Exame	Esperado	RNA	Exame	Esperado	RNA
1	0	0,026300	13	0	0,086500
2	0	0,000641	14	0	0,003700
3	0	0,000005	15	0	0,012400
4	0	0,000054	16	1	0,948600
5	0	0,000398	17	0	0,000366
6	0	0,019500	18	0	0,005900
7	0	0,000367	19	0	0,000005
8	0	0,000041	20	0	0,000105
9	1	0,999700	21	0	0,002500
10	0	0,096400	22	0	0,019400
11	0	0,000025	23	0	0,000004
12	0	0,000000			

Fonte: Autoria própria, 2023.

Na Tabela 6 é mostrado o resultado obtido pela RNA de 100 neurônios, para os exames que os profissionais da medicina avaliaram de forma diferente. De acordo com essa RNA, o primeiro exame teria 29 % de chance de se ter uma infecção

bacteriana, que se aproximaria das avaliações feitas pelo médico e pela enfermeira 2. O segundo exame seria 90% de chance de apresentar infecção bacteriana, novamente semelhante as avaliações feitas pelo médico e pela enfermeira 2.

Tabela 6. RNA com 100 neurônios x Profissionais da medicina

Exame	RNA	Médico	Enfermeira 1	Enfermeira 2
1	0,295783	0	1	0
2	0,906933	1	0	1
3	0,385357	1	0	1
4	0,477139	1	0	0
5	0,068377	1	0	0
6	0,558196	1	0	0
7	0,999936	1	1	0
8	0,169355	1	0	0

Fonte: Autoria própria, 2023.

O exame 3, com 38% de chance de infecção bacteriana, se assemelharia mais com a opinião da enfermeira 1. Observando o exame 4, com 47 % de chance de ser uma infecção bacteriana, auxiliaria no caso de quem deu avaliação negativa de infecção bacteriana, gerando a possibilidade de se pedir um outro exame ou olhar esse mesmo exame com um pouco mais de cuidado.

Analisando o exame 5 com 6% de chance de infecção bacteriana, descartaria assim a chance de infecção bacteriana, sendo semelhante as duas enfermeiras. No hemograma 6 com 55 % de possibilidade de ser uma infecção, deveria ser realizada uma nova avaliação dos exames feitos pelas enfermeiras.

Fazendo a análise do hemograma 7 com 99% de possibilidade de infecção bacteriana, a rede foi semelhante as avaliações do médico e da enfermeira 1 e diferente da enfermeira 2. Finalmente a última avaliação, o exame 8 com 16% de chance de infecção bacteriana, ajudaria o profissional da saúde a rever sua interpretação para que se pudesse evitar receitar antibióticos sem necessidade.

A Tabela 7 apresenta o resultado dos exames divergentes obtidos pela RNA que possui 200 neurônios na camada de entrada.

Tabela 7. RNA com 200 neurônios x Profissionais da medicina

Exame	RNA	Médico	Enfermeira 1	Enfermeira 2
1	0,120402	0	1	0
2	0,000212	1	0	1
3	0,999713	1	0	1
4	0,040869	1	0	0
5	0,032972	1	0	0
6	0,086604	1	0	0
7	0,650980	1	1	0
8	0,002755	1	0	0

Fonte: Autoria própria, 2023.

No hemograma 1, com 12 % de possibilidade de infecção bacteriana, a rede teve um comportamento avaliativo semelhante ao médico e a enfermeira 2, o que aconteceu com a rede anterior. Avaliando o resultado do exame 2, com 0 % de chance de infecção bacteriana, a avaliação se assemelhou a da enfermeira 1.

Continuando com as análises, observando o exame 3, com 99% de chance de infecção bacteriana, a rede reforçaria as avaliações realizadas pelo médico e pela enfermeira 2, e iria propor uma revisão pela enfermeira 1. No hemograma 4 com 4% de possibilidade de infecção bacteriana, ajudaria a enfermeira 1 e 2 terem certeza de suas avaliações e ao médico rever o que foi avaliado.

Realizando-se a observação do resultado do hemograma 5 feito pela RNA, apresentando 3 % de possibilidade de uma infecção por bactérias, novamente iria ajudar a reforçar as análises das enfermeiras e uma reavaliação pelo médico. Analisando o exame 6 com 8% de chance de infecção bacteriana, realçaria o que foi avaliado pelas enfermeiras e abriria a oportunidade de se observar o exame com mais cuidado pelo médico.

Fazendo a análise do exame 7 segundo o que a RNA apresentou como sugestão, com 65,09% de chance de ser uma infecção bacteriana, reforçaria as interpretações que o médico e a enfermeira 1 fizeram e auxiliaria a enfermeira 2 a observar novamente sua interpretação. Por último, o hemograma 8 com 0,2 % de possibilidade de ser uma infecção bacteriana, reforçaria o que foi interpretado pelas enfermeiras e iria ajudar o médico que possivelmente poderia mudar o diagnóstico.

Finalmente, analisando os resultados obtidos pela RNA com 500 neurônios, mostrada na Tabela 8.

Tabela 8. RNA com 500 neurônios x Profissionais da medicina

Exame	RNA	Médico	Enfermeira 1	Enfermeira 2
1	0,994995	0	1	0
2	0,841023	1	0	1
3	0,468355	1	0	1
4	0,778682	1	0	0
5	0,092451	1	0	0
6	0,630618	1	0	0
7	0,875211	1	1	0
8	0,160008	1	0	0

Fonte: Autoria própria, 2023.

Observa-se que o exame 1 com cerca de 99% de chance de infecção bacteriana, se assemelha a opinião dada pela enfermeira 1, diferente do que acontece com as redes anteriores. Interpretando o hemograma 2 com 84 % de chance de infecção bacteriana, reforçaria o que foi analisado pelo médico e pela enfermeira 2.

Realizando a análise do exame 3, com 47 % de possibilidade de ser uma infecção bacteriana, a sugestão de resultado da rede ajudaria tanto um lado como o outro a reverem suas interpretações, principalmente a enfermeira 1, que havia considerado que este não apresentava infecção. Para o hemograma 4, com 77 % de chance de infecção bacteriana, reforçaria o que o médico observou e ajudaria as enfermeiras a analisarem novamente e talvez mudarem sua interpretação.

Observando-se o hemograma 5 com 9 % de chance de infecção, percebe-se que a sugestão da rede ajudaria o médico a analisar novamente a sua interpretação. Com a análise obtida pelo exame 6 percebe-se que com 63% de possibilidade de infecção bacteriana, reforça a interpretação do médico.

Analisando o exame 7 que apresentou 87% de possibilidade da doença, ajudaria a enfermeira 2 a rever seu laudo. Finalmente, o exame 8 com 17 % de chance de infecção bacteriana, reforçaria as interpretações que as enfermeiras tiveram.

Portanto, analisando os resultados obtidos pelas 3 Redes Neurais, a rede neural com 500 neurônios pode ser considerada a melhor das 3 treinadas, tendo em vista que, essa quantidade de neurônios se aproxima do limite de generalização que essa rede permite.

Capítulo 5

5 Conclusões e trabalhos futuros

Através dos resultados obtidos, pode-se perceber que a Rede Neural é uma excelente ferramenta para auxílio na interpretação de hemogramas e para detecção de infecções bacterianas, tendo em vista os resultados obtidos pela mesma nos treinamentos realizados, levando-se em consideração o tamanho da amostra utilizada e os excelentes resultados obtidos com cerca de 96% e 100% de acurácia frente os 85% que é a média dos desempenhos de Redes Neurais encontrada na literatura, segundo Veronezi et al. (2011). Esses 96% a 100% acurácias foram obtidos se considerarmos que a rede com 100 neurônios teve uma análise diferente do que era esperado no primeiro teste, e as demais tiveram análises iguais.

O percentual que as RNAs desempenharam também é ótimo, se considerarmos as discordâncias diagnósticas na avaliação do total de exames. Em cerca de 93 % dos exames houve concordância nos diagnósticos, e 7 % de discordância. Por outro lado, com essa taxa de acerto variando de 96 % a 100 %, as RNAs treinadas tiveram um ótimo desempenho, demonstrando ser uma poderosa ferramenta auxiliadora, tendo em vista que ela se comporta como a junção das opiniões dos 3 avaliadores.

Para trabalhos futuros, sugere-se que se possa trabalhar com uma quantidade maior de amostras para se ter uma validação maior, além de ser possível que ela seja mais homogênea na quantidade de casos negativos e positivos, na distribuição quantitativa de sexo dos pacientes e uma boa uniformidade na quantidade de exames por faixa etária. Pois assim, a rede teria um desempenho melhor e passaria uma maior confiança em relação ao seu desempenho.

Outra sugestão é ter mais avaliadores além dos 3 utilizados, que funcionariam como uma junta médica. Consequentemente isso talvez resultaria em um número maior de exames com avaliações em que esses profissionais discordariam.

Aumentando o percentual da amostra total usado para o teste 2, no qual já seria avaliado a interpretação desses exames pela Rede Neural.

Outra possibilidade seria classificar mais tipos de infecções, além da bacteriana que é objeto de estudo dessa dissertação. Levando-se em consideração também os limites do banco de dados oferecidos pelo hospital, que diante da interpretação médica apresentavam só infecções bacterianas em uma quantidade suficiente para treinamento da rede.

Poderia ser feita a classificação dos exames nos diferentes tipos de infecções, caso possível, ou somente dar mais um passo à frente e ter infecções bacterianas e infecções virais. Até mesmo, caso possível, dividir as infecções bacterianas nos diversos tipos apresentados, mas para isso necessitaria de um banco de dados maior e que apresentasse várias amostras dos diferentes tipos de infecções que forem estudadas.

Referências

BANDEIRA, R., MAGALHÃES, A. F., DE AQUINO, H. B. da S. Interpretação dos critérios de liberação dos resultados de hemograma através de contadores automatizados em laboratórios de urgência. **Revista Saúde e Pesquisa**, Maringá, v. 7, n.3, 2014.

BRAZÃO-SILVA, M. T. **Eritrograma para prática odontológica: Revisão de literatura**, p. 146-154, 2019.

CARVALHO, R. de. **Revolução do Hemograma. Aprenda a interpretar e relacionar marcadores para ter um excelente diagnóstico clínico**. Disponível em: <<https://gabrieldecarvalho.com.br/>> Acesso em: 24/03/2021.

CARVALHO, R. D. de; MACEDO, L. P.; TEIXEIRA, F. A.; BINDA, M. B.; COELHO, C. S. Volume corpuscular médio (VCM) e amplitude da distribuição do tamanho dos eritrócitos (RDW) em equinos da raça quarto de milha usados em provas de três tambores. **Ciência Animal Brasileira**, Goiânia, v. 17, n. 3, p. 411-417, 2016.

CHAUFFAILLE, M. de L. L. F. Eosinofilia reacional, leucemia eosinofílica crônica e síndrome hipereosinofílica idiopática. **Revista Brasileira de Hematologia e Hemoterapia**, São Paulo, v. 32, n. 5, p. 395-401, 2010.

COSTA, F.; MOTTA, L.; NOGUEIRA, J. Uma abordagem baseada em Redes Neurais Artificiais para o auxílio ao diagnóstico de doenças meningocócicas. **Revista Brasileira de Computação Aplicada**, v. 2, n. 1, p. 79-88, 2010.

DE ALMEIDA, M.; DE MELO, C.; CÂMARA DE ALMEIDA, M. Causas de Alterações morfológicas nos glóbulos vermelhos que comprometem o resultado do laudo clínico. **Revista de Ciências da Saúde Nova Esperança**, v. 10, n.1, p. 84-91, 15 jun. 2012.

EHRLICH, P. Beiträge zur Kenntnis der granulierten Bindegewebszellen und der eosinophilen Leukocythen. **Arch Anat Physiol**, Leipzig, v. 3, p. 166–169, 1879.

FAILACE, R.; FERNANDES, F. **Hemograma: manual de interpretação**. 6. ed. – Porto Alegre: Artmed, 2015.

FALCON DÍAZ, M.; CARMONA, E. G.; VALDÉS, D. V.; PALACIOS, Y. T.; RODRÍGUEZ, Y. S. The diagnostic importance of the leukogram in the febrile and post-

febrile stage in adults. **Revista Ciências Médicas**, Pinar del Río, v. 13, n. 2, p. 169-184, 2009.

GOULART, D. A.; TACLA, M. A.; MARBACK, P. M. F.; SOLÉ, D.; JUNIOR, A. P.; PEREZ, H. B.; DE FREITAS, D.; SATO, E. H. Redes neurais artificiais aplicadas no estudo de questionário de varredura para conjuntivite alérgica em escolares. **Arquivos Brasileiros de Oftalmologia**, 69(5), 707-713, 2006.

HELEGDA, L. C.; HELEGDA, S.; PETRACCO, A. M.; BODANESE, L. C. Predição à doença cardiovascular utilizando redes neurais artificiais incluindo níveis séricos de IL-6 em pacientes com síndrome metabólica. **Revista Ciência e Tecnologia**. V.15. n. 2, p 14-23, 2015.

LOBO, L. X. Inteligência artificial e medicina. **Revista Brasileira de educação médica**, Brasília, 2017.

MCCULLOCH, W.S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**. Vol 5, n. 5, pp 115-133, 1943.

PEREIRA, J. L. **Identificação de danos estruturais baseada em um modelo de dano contínuo via redes neurais artificiais**. Dissertação de mestrado, Universidade Federal Fluminense, Volta Redonda, 2018.

RAMOS, A. C. S.; SOUZA I. C. L. Alterações hematológicas em pacientes com infecções bacterianas. **Anais, 18ª Semana de Pesquisa Universidade Tiradentes**. “A prática interdisciplinar alimentando a ciência.”. 2016.

RAMOS, M. A. F. **Sistema de Identificação e Localização de Faltas em Linhas de Transmissão Baseado em Redes Neurais**, 2002.

ROSENFELD, R. Hemograma. **Jornal Brasileiro de Patologia e Medicina laboratorial**, Rio de Janeiro, v. 48, n. 4, p. 244, 2012.

SALOMÃO, R. **Infectologia, Bases clínicas e tratamento / Reinaldo Salomão – 1. ed.** – Rio de Janeiro: Guanabara Koogan, 2017.

SANTOS V. S. dos. “O que é infecção?”; **Brasil Escola**. Disponível em: <https://brasilecola.uol.com.br/o-que-e/biologia/o-que-e-infeccao.htm>. Acesso em 12 de abril de 2021.

SANTOS, N. de Q. A resistência bacteriana no contexto da infecção hospitalar. **Texto & Contexto – Enfermagem**, v. 13, n. spe, p. 64-70, 2014.

SILVA, I. N. D.; SPATTI, H.; FLAUZINO, R. **Redes neurais artificiais, curso prático para engenharia e ciências aplicadas**. Artliber Editora, 2010.

SILVA, I. N. D.; SPATTI, H.; FLAUZINO, R. **Redes neurais artificiais, curso prático para engenharia e ciências aplicadas**. Artliber Editora, 2016.

SIQUEIRA-BATISTA, R.; VITORINO, R. R.; GOMES, A. P.; OLIVEIRA, A. DE P.; FERREIRA, R. DOS S; ESPERIDIÃO-ANTONIO, V.; SANTANA, L. A.; CERQUEIRA, F. R. As redes neurais artificiais e o ensino da medicina. **Revista Brasileira de Educação Médica**, Rio de Janeiro, v. 38, n. 4, p. 548-556, 2014.

VERONEZI, C. C. D.; SIMÕES, P. W. T. DE A.; SANTOS, R. L. DOS.; DA ROCHA, E. L.; MELÃO, S.; DE MATTOS, M. C.; CECHINEL, C. Análise computacional para auxílio ao diagnóstico de osteoartrite de coluna lombar baseado em redes neurais artificiais. **Revista Brasileira de Ortopedia**, São Paulo, v. 46, n.2, p. 195-199, 2011.

WINTROBE, Maxwell M. Variations in the size and hemoglobin content of erythrocytes in the blood of various vertebrates. **Folia haematologica**, v.51, n.32, p. 32.49,1934.