



MÉTODOS ESTATÍSTICOS E NEURAIIS PARA DIAGNÓSTICO DA  
TUBERCULOSE BOVINA VIA TESTES CERVICAIS.

Fellipe de Oliveira Pinto

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Luiz Pereira Calôba

Rio de Janeiro

Maio de 2015

MÉTODOS ESTATÍSTICOS E NEURAIIS PARA DIAGNÓSTICO DA  
TUBERCULOSE BOVINA VIA TESTES CERVICAIS.

Fellipe de Oliveira Pinto

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO  
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE  
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE  
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A  
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA  
ELÉTRICA.

Examinada por:

---

Prof. Rubens de Andrade Junior, D.Sc.

---

Prof. Walter Lilenbaum, D.Sc.

---

Prof. José Gabriel Rodriguez Carneiro Gomes, Ph.D.

RIO DE JANEIRO, RJ – BRASIL  
MAIO DE 2015

Pinto, Fellipe de Oliveira

Métodos Estatísticos e Neurais para Diagnóstico da Tuberculose Bovina via Testes Cervicais./Fellipe de Oliveira Pinto. – Rio de Janeiro: UFRJ/COPPE, 2015.

XIII, 77 p.: il.; 29, 7cm.

Orientador: Luiz Pereira Calôba

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2015.

Referências Bibliográficas: p. 73 – 77.

1. Redes Neurais Artificiais. 2. Tuberculose Bovina. 3. Classificação Supervisionada. I. Calôba, Luiz Pereira. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*Aos meus pais pelo dom da vida  
e pelo amparo ao longo desses  
anos.*

*“Tudo posso Naquele que me  
conforta”. Filipenses,4:13*

# Agradecimentos

Inicio meus agradecimentos a DEUS por iluminar o meu caminho, por me dar uma família maravilhosa que sempre me apoiou em tudo que pretendi e por me dar amigos tão importantes na minha vida.

Aos meus pais, José e Maria, meu infinito agradecimento. Sempre acreditaram em minha capacidade e me incentivaram a persistir nos estudos. Obrigado pelo vosso amor incondicional!

A minha irmã Fernanda, pois, a seu modo, sempre se orgulhou de mim e confiou em meu trabalho.

Aos meus “irmãos” Fábio e Priscila que sempre me ajudaram a tomar as decisões certas sempre com muito carinho e atenção, sem a presença deles com certeza seria muito mais difícil.

Ao meu orientador Luiz Pereira Calôba por aceitar a me orientar e sempre estar disponível para me auxiliar em todos os aspectos desta dissertação; sem seu auxílio acredito que não seria possível.

Ao Professor José Gabriel Rodriguez Carneiro Gomes por aceitar pertencer a banca examinadora e por ter sido meu professor durante o curso de mestrado

Ao Professor Walter Lilenbaum que além de aceitar pertencer a banca Examinadora me auxiliou com a base de dados e com o conhecimento sobre tuberculose bovina.

Ao Professor Rubens de Andrade Júnior por aceitar participar da banca examinadora

A minha maior incentivadora depois da família, a Professora Regina Serrão Lanzillotti sem ela talvez nunca tivesse pensado em fazer mestrado, e sua irmã Haydé Lanzillotti por também me incentivar.

A todos os demais professores que, de certo modo, contribuíram para meu desenvolvimento pessoal e profissional em especial os Professores José Francisco e Mariane Branco.

A minha amiga Roberta Pereira Niquini que sempre acreditou em mim e me incentivou em todos os momentos.

Ao meu amigo João Paulo Zanardi sempre disponível e disposto para ajudar sempre.

A minha amiga Michelle Fraga por realizar a revisão gramatical e por ser minha amiga.

Ao meus amigos Renata Monteiro e Felipe Niquini por além de serem meus amigos fiéis, me convidaram a ser padrinhos de seu casamento, notícia essa que me alegrou muito.

Aos meus amigos Alexandre e Lidiane que também me deram a honra de ser padrinho de vosso casamento e por, me tratem como se fosse de sua família.

Ao meu amigo Márcio Brandão que sempre me incentivou a fazer o mestrado.

Ao amigo Savano que além de estudar comigo sempre teve paciência e também pelas caronas que fizeram eu economizar tempo nas idas ao Fundão.

A minha amiga Aline Botelho por sempre me apoiar, me ouvi e rezar por mim!

Aos meus afilhados Gabriel e Sofia. além é claro do meu sobrinho Guilherme que me proporcionaram momentos de felicidade ao longo desta jornada.

A minha família, tios, tias, primos e primas, em especial aos primos Rodrigo, Beatriz e Diego pelo carinho e união nos momentos de felicidade e tristeza.

Aos Demais amigos que fiz no curso de Estatística: Bruno, Carol, Diogo, Juliana, Vitor, Aline, Marcio Santos, Marcelo Goulart, Marília, Flávia, Monique, Laura, Roberto e Felipe Rafael pela amizade sincera!

Aos demais amigos que fiz no curso de matemática: Luis Phelipe, Bruno Paiva, Isabela, Thiago José, Fernanda, Claudio, Vitor, Alex, Felipe Menezes, Sabrina Mota e André Luiz pela amizade sincera!

Aos amigos que fiz no Mestrado André Luiz, Igor e Natanael Junior por auxiliarem em minha dissertação.

Agradeço, também, ao CNPq pelo apoio financeiro.

Gostaria de agradecer aos Funcionários do Programa de Engenharia Elétrica que sempre me atenderam com atenção, em especial a Daniele.

Por fim, gostaria de agradecer ao Programa de Engenharia Elétrica por abrir as portas para que eu pudesse realizar este sonho. Proporcionaram-me mais que uma busca pelo conhecimento técnico e científico, uma lição de vida.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

## MÉTODOS ESTATÍSTICOS E NEURAI PARA DIAGNÓSTICO DA TUBERCULOSE BOVINA VIA TESTES CERVICAIS.

Fellipe de Oliveira Pinto

Maio/2015

Orientador: Luiz Pereira Calôba

Programa: Engenharia Elétrica

A tuberculose bovina (TBb) é uma zoonose cujo hospedeiro primário é o bovino, mas que também pode infectar o homem, e por isto a TBb é considerada uma doença de alta relevância social e econômica. O exame preconizado para seu diagnóstico são os teste tuberculínicos, que entretanto apresentam sérios problemas de sensibilidade e especificidade. Estes problemas não permitem uma acurácia satisfatória no diagnósticos da TBb. Além disto, a permanência no rebanho de animais infectados leva a ocorrência de novos casos dentro do próprio rebanho, tornando-se assim um obstáculo para o sucesso do PNCEBT. Neste trabalho, estudaram-se a utilização de modelos discriminantes e neurais na tentativa de aumento da acurácia dos diagnósticos de tuberculoses em bovino. Foram determinados também, novos pontos de corte para os testes cervicais e desenvolvidos modelos neurais com  $EQM_{modificado}$  e modelos neurais com polarização positiva com o objetivo de aumentar a sensibilidade do diagnóstico sem que houvesse uma redução drástica da especificidade. O estudo evidencia que o problema é de difícil solução e que, os modelos discriminantes e neurais clássicos foram ineficientes em aumentar a acurácia dos diagnósticos. Em relação aos procedimentos realizados para aumento da sensibilidade o modelo com  $EQM_{modificado}$  se mostrou mais habilitado a maximizar a sensibilidade, uma vez que a redução da especificidade ocorreu de maneira menos severa que nos pontos de corte dos testes cervicais. Portanto, julga-se o modelo neural com  $EQM_{modificado}$  promissor no diagnóstico com sensibilidade máxima e, acredita-se que este trabalho pôde ter contribuído para a pesquisa e desenvolvimento do controle da Tuberculose bovina.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

STATISTICAL METHODS AND NEURAL NETWORKS BOVINE  
TUBERCULOSIS DIAGNOSIS USING CERVICAL TESTS

Fellipe de Oliveira Pinto

May/2015

Advisor: Luiz Pereira Calôba

Department: Electrical Engineering

Bovine tuberculosis (bTB) is a zoonosis whose primary host is cattle, but can also infect humans, and therefore the bTB is considered a high social and economic relevance disease. The recommended test for its diagnosis are the tuberculin tests, which however present serious problems of sensitivity and specificity. These problems do not allow a satisfactory accuracy in the diagnosis of bTB. Furthermore, the persistence in infected bovine leads to occurrence of new cases within the herd itself, thus becoming an obstacle to the success of PNCEBT. In this work, we estimated discriminants and neural models in an attempt to increase the accuracy of diagnosis of tuberculosis in cattle. Were determined also, new breakpoints for cervical tests and developed neural models with  $MSE_{modified}$  and neural models with positive bias in order to increase the sensitivity of diagnosis without there was a drastic reduction in specificity. The study shows that the problem is difficult to solve and that the discriminants and neural models were unable in increasing the accuracy of diagnosis. Regarding the procedures performed to increase sensitivity, the model with  $MSE_{modified}$  more able to maximize sensitivity, since the reduction in specificity was less severe so that the cutting points was defined the new rule tuberculin tests. Therefore, it is thought the neural model with  $MSE_{modified}$  promising in the diagnosis with maximum sensitivity and, it is believed that this work could have contributed to the research and development of the control of bovine tuberculosis.



# Sumário

<b>Agradecimentos</b>	<b>v</b>
<b>Lista de Figuras</b>	<b>xi</b>
<b>Lista de Tabelas</b>	<b>xii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Visão Geral . . . . .	1
1.2 Motivação . . . . .	2
1.3 Objetivo . . . . .	3
1.4 Organização da dissertação . . . . .	4
<b>2 Tuberculose Bovina</b>	<b>5</b>
2.1 Epidemiologia . . . . .	5
2.2 Diagnóstico . . . . .	6
2.3 Diagnósticos diferenciais . . . . .	9
2.4 Prognóstico . . . . .	9
2.5 Controle e Profilaxia . . . . .	10
2.6 Impacto Econômico da TB bovina . . . . .	11
<b>3 Métodos de classificação Supervisionada</b>	<b>12</b>
3.1 Métodos Estatísticos de Classificação Supervisionada . . . . .	13
3.1.1 Análise Discriminante . . . . .	13
3.1.2 Função Discriminante Linear de Fisher . . . . .	15
3.1.3 Função Discriminante Quadrática . . . . .	16
3.2 Redes Neurais . . . . .	17
3.2.1 Neurônio Biológico . . . . .	17
3.2.2 Neurônio Artificial . . . . .	18
3.2.3 Características de uma Rede Neural Artificial . . . . .	21
3.2.4 Redes Multicamadas <i>Perceptron</i> (MLP) . . . . .	23
3.2.5 Algoritmo <i>backpropagation</i> . . . . .	26
3.2.6 Algoritmo <i>Resilient Propagation</i> . . . . .	31

3.3	Métodos de Avaliação dos modelos Supervisionados . . . . .	33
3.3.1	Curva ROC . . . . .	34
3.3.2	Validação Cruzada . . . . .	35
<b>4</b>	<b>Metodologia</b>	<b>37</b>
<b>5</b>	<b>Resultados</b>	<b>45</b>
5.1	Análise dos Testes Cervicais . . . . .	46
5.2	Modelos Discriminantes . . . . .	50
5.3	Resultados da rede neural clássica . . . . .	54
5.4	Resultados da rede neural com Erro Quadrático variável . . . . .	56
5.5	Resultados da rede neural com polarização diferenciada . . . . .	60
5.6	Comparação entre os Testes cervicais e os principais modelos . . . . .	61
<b>6</b>	<b>Conclusão</b>	<b>64</b>
6.1	Discussão . . . . .	64
6.2	Considerações Finais . . . . .	71
	<b>Referências Bibliográficas</b>	<b>73</b>

# Lista de Figuras

3.1	Representação de um Neurônio Biológico. . . . .	18
3.2	Modelo matemático de um neurônio biológico (McCulloch & Pitts [41])	19
3.3	Modelo de neurônio base para projetos de redes neurais artificiais . .	20
3.4	RNA de estrutura simplificada <i>feedforward</i> e com uma camada escondida e outra de saída. . . . .	23
3.5	Validade de um teste de diagnóstico ou dos resultados de um modelo de classificação. . . . .	34
5.1	Plano $\Delta B_{ov}$ versus $\Delta B_{ov} - \Delta A_v$ . . . . .	45
5.2	Curva ROC dos testes de diagnósticos $\Delta B_{ov}$ e $\Delta B_{ov} - \Delta A_v$ . . . . .	47
5.3	Separador Discriminante Linear (a) e Discriminante Quadrático (b) .	52
5.4	Saídas da Rede neural com 3 neurônios . . . . .	55
5.5	Representação dos separadores das redes neurais projetadas para cada valor do parâmetro $M$ . . . . .	59
5.6	Curvas ROC das variáveis $\Delta B_{ov}$ e $\Delta B_{ov} - \Delta A_v$ e dos modelo discriminante linear e modelo neural clássico com 3 neurônios na camada intermediária . . . . .	62
5.7	Curvas ROC das variáveis $\Delta B_{ov}$ e $\Delta B_{ov} - \Delta A_v$ , modelo neural modificado por $M = 4$ e modelo neural clássico com polarização positiva de $W = 0, 14$ . . . . .	63
5.8	Curvas ROC das variáveis $\Delta B_{ov}$ e $\Delta B_{ov} - \Delta A_v$ , modelo neural modificado por $M = 3, 5$ e modelo neural clássico com polarização positiva de $W = 0, 12$ . . . . .	63

# Lista de Tabelas

2.1	Interpretação do Teste Cervical Simples em bovinos [7] . . . . .	8
2.2	Interpretação do teste cervical comparativo em bovinos [7] . . . . .	9
4.1	Critério de classificação dos resultados do teste cervical simples . . .	37
4.2	Critério de Classificação dos Resultados do Teste Cervical Comparativo	38
4.3	Valores do Parâmetro $M$ utilizados na ponderação da função indica- dora 4.1 . . . . .	42
4.4	Tabela de variações $W$ na polarização da rede neural com parâmetro $M = 1$ . . . . .	44
5.1	Resultados dos Testes de diagnósticos TCS e TCC quanto à Tubercu- lose Bovina . . . . .	46
5.2	Índices de desempenho dos Testes de diagnósticos considerando os resultados de bovinos inconclusivos como infectado por Tuberculose .	47
5.3	Pontos de corte para a variável $\Delta Bov$ e seus respectivos índices de sensibilidade, falso alarme e SP . . . . .	49
5.4	Pontos de corte para para a variável $\Delta Bov - \Delta Av$ e seus respectivos índices de sensibilidade, falso alarme e SP . . . . .	50
5.5	Resultados dos teste de normalidade Komolgorov-Smirnov para as variáveis independentes . . . . .	51
5.6	Resultados do teste Box's M de homogeneidade das matrizes de co- variâncias das variáveis independentes . . . . .	51
5.7	Índices de desempenho na classificação de Tuberculose Bovina pelos Modelos Discriminante Linear e Quadrático para os conjuntos treino e teste . . . . .	53
5.8	Validação Cruzada dos modelos discriminante Linear e Quadrático . .	54
5.9	Resultados dos Projetos de Rede neurais utilizando topologias diferentes	55
5.10	Validação Cruzada para Rede neural que utiliza 3 neurônios na ca- mada intermediária . . . . .	56

5.11	Resultados dos modelos neurais utilizando 3 neurônios na camada intermediária e com as ponderações pelos valores de M na função objetivo . . . . .	60
5.12	Validação Cruzada para rede neural que utiliza 3 neurônios na camada intermediária e $EQM_{modificado}$ com parâmetro $M=3,5$ . . . . .	60
5.13	Resultados das polarizações <i>bias</i> nos modelos neurais utilizando 3 neurônios na camada intermediária e sem ponderação na função objetivo	61

# Capítulo 1

## Introdução

### 1.1 Visão Geral

A tuberculose bovina (TB bovina) é uma zoonose, doença transmissível entre os animais vertebrados e o homem, causada por *Mycobacterium bovis* (*M. bovis*), cujo hospedeiro primário é o bovino. Além do homem, diversas outras espécies de mamíferos domésticos e silvestres são também susceptíveis ao mesmo bacilo [4], [15].

A TB bovina pode ser caracterizada pela formação de lesões do tipo granulomatosa, por ser uma doença crônica e por apresentar os sintomas tardiamente, apenas quando há comprometimento dos órgãos atingidos [29] [1].

Essa zoonose apresenta uma significativa prevalência nos países em desenvolvimento. Por outro lado, em países desenvolvidos, a tuberculose bovina possui prevalência baixa ou, até mesmo, sua completa erradicação, resultados que são consequência da implantação e consolidação de programas de controle e erradicação[35].

Nesse Contexto, o Ministério da Agricultura, Pecuária e Abastecimento instituiu, em 2001, o Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose (PNCEBT), cujo principal objetivo era diminuir o impacto negativo dessas zoonoses na saúde pública humana e animal [49], além é claro, de diminuir as barreiras sanitárias que são impostas pelos principais mercados consumidores[7] [56][18]. Também é conhecido que o exame preconizado pelo PNCEBT que é utilizado para certificar a sanidade do rebanho é muitas vezes limitado, assim, é de extrema importância para o sucesso do PNCEBT que se busque aperfeiçoar o processo de diagnóstico da tuberculose bovina para que este seja o mais preciso possível e assegure uma rápida e sensata decisão [58].

Segundo Friedman “*um processo de diagnóstico pode ser visto como uma tarefa de decisão realizada com base nos sinais clínicos, nos sintomas e outros indícios laboratoriais e, frequentemente envolve incertezas devido às variações entre os pacientes, erros de aferições ou ausências de observações*” [26].

Utiliza-se testes diagnósticos como instrumento auxiliar em um processo de diagnóstico. O uso destes testes vem se tornando bastante comum e, em geral, são avaliados a partir de medições feitas no paciente, sejam elas quantitativas ou qualitativas [26]. Ao se avaliar o resultado de um teste diagnóstico, quatro situações são possíveis: o teste é positivo, mas o paciente não tem a doença (falso positivo); o teste é negativo e o paciente não tem a doença (verdadeiro negativo); o teste é positivo, e o paciente tem a doença (verdadeiro positivo) e o teste é negativo e o paciente tem a doença (falso negativo)[26].

Obviamente, os testes de maior utilidade clínica são aqueles que apresentam poucos resultados falsos positivos e falsos negativos, porém, na prática, nem sempre esta realidade é possível, o que torna o processo de diagnóstico muito impreciso se for utilizado apenas um teste de diagnóstico[26].

Mediante esse problema e com o desenvolvimento de algoritmos facilitado pelo avanço a técnicas computacionais, o processo de diagnóstico muitas vezes tem sido aprimorado através da combinação dos resultados de diversos testes diagnósticos por meios de modelos estatísticos, análise discriminante, ou de inteligência artificial, como redes neurais[20].

A Análise Discriminante constitui-se em um modelo estatístico com o propósito de construir regras de decisão que vão alocar um novo indivíduo  $X$  a uma das  $g$  classes distintas, previamente conhecidas, admitindo-se que  $X$  realmente pertença a uma dessas  $g$  classes. O tratamento estatístico dado a esse problema de locação reside no fato que os dados utilizados são os valores de um conjunto de  $p$  variáveis aleatórias  $(X_1, \dots, X_p)$  e serão construídas funções discriminantes a partir delas [32].

As Redes Neurais Artificiais são modelos computacionais e matemáticos inspirados no cérebro capazes de resolver problemas não-lineares complexos de aproximação, incluindo o reconhecimento de padrões [31].

Existem diversas estruturas de redes neurais, mas de uma forma geral, uma rede neural artificial é composta por unidades de processamento que são denominados neurônios conectados por pesos sinápticos, parâmetros do modelo que são ajustados por um algoritmo iterativo e paralelo através dos dados. Uma vez ajustados os pesos, a rede tem a capacidade de aprender através dos dados amostrais e de generalizar para novos casos [31].

## 1.2 Motivação

Há diversos aspectos relacionados à motivação deste trabalho.

O primeiro aspecto é a relevância econômica da tuberculose bovina, uma vez que causa danos aos sistemas produtivos, influenciando nos custos de produção, principalmente na reposição de animais eliminados quando positivos aos testes de

diagnóstico preconizados pelo PNCEBT, como recomenda a legislação. Além, é claro, da influência no valor agregado dos animais que é reduzido juntamente com o dos seus subprodutos no comércio internacional [43].

O segundo aspecto é em relação à saúde pública veterinária, pois ainda existe a necessidade de melhorias na prevenção da infecção por *M.bovis*, especialmente nas populações humanas em risco, como os tratadores de rebanhos e trabalhadores da indústria de carnes e laticínios [8] [48].

O terceiro aspecto que motivou este trabalho foi a existência de limitações do método de diagnóstico preconizado pelo PNCEBT, que utilizava, muitas vezes, seu uso como ferramenta única de diagnóstico não permitindo a detecção de todos os animais infectados, tornando-se um dos principais obstáculos para o sucesso do controle da tuberculose bovina [37].

Por fim, acredita-se na potencialidade da utilização de modelos que possam combinar diversos testes de diagnósticos, pois estes modelos tem sido muito utilizados na área médica devido ao aprimoramento da multidisciplinaridade destas áreas do conhecimento.

### 1.3 Objetivo

Diante do exposto na seção anterior, um dos objetivos deste trabalho é promover o desenvolvimento de modelos estatísticos e neurais que ao combinarem os resultados dos testes de diagnósticos preconizados pelo PNCEBT podem auxiliar no processo de diagnóstico da tuberculose bovina com o aumento da acurácia.

Além disso, o outro objetivo tem por finalidade elevar o poder de detecção de bovinos infectados com TB bovina implementando modelos que aumentem a sensibilidade do diagnóstico sem que haja uma redução drástica da especificidade.

Ainda espera-se, através dos resultados obtidos, poder realizar avaliações entre os testes de diagnósticos preconizados pela legislação, os modelos estatísticos e as redes neurais artificiais clássicas, a fim de estabelecer o modelo mais indicado para o aumento da acurácia dos diagnósticos.

De forma análoga, foram avaliados os métodos aplicados com a finalidade de aumentar a sensibilidade do diagnóstico (novas regras de decisão dos Testes cervicais, redes neurais com modificações e redes neurais com polarização positiva).

Desta maneira, foi possível ponderar as vantagens e desvantagens de cada método para o diagnóstico da tuberculose bovina.



## 1.4 Organização da dissertação

Além desta introdução, o presente trabalho contém mais cinco capítulos, organizados da forma descrita a seguir.

No capítulo 2 é feita uma revisão bibliográfica sobre tuberculose bovina, onde é apresentado uma descrição sobre a epidemiologia, diagnósticos, prognósticos e impacto econômico no mercado exterior.

O capítulo 3 aborda os conceitos básicos dos modelos discriminantes e de redes neurais, a contextualização dos principais pressupostos na modelagem estatística, uma descrição do modelo discriminante linear e do modelo quadrático, bem como, características de redes neurais, tipo de função ativação e a descrição dos Algoritmos *Backpropagation* e *Resilient Propagation*.

Os capítulos 4 e 5, por sua vez, representam o cerne prático deste trabalho. No capítulo 4 é demonstrado o passo a passo da execução das análises propostas: avaliação dos testes preconizados pelo PNCBT, propostas de novo ponto de corte através do estudo da curva ROC, modelos estatísticos e projetos de redes neurais. No capítulo 5 são analisados cada um dos resultados obtidos, bem como comentada a comparação entre as técnicas investigadas.

Finalizando, o capítulo 6 traz uma discussão sobre os principais métodos de diagnósticos alternativos, além de uma avaliação dos resultados encontrados e uma conclusão sobre o estudo realizado com os avanços alcançados, além de propostas de novos estudos.

# Capítulo 2

## Tuberculose Bovina

A tuberculose bovina (TB bovina), é uma doença infecciosa e de ocorrência mais provável em mamíferos, incluindo o homem [17]. Assim, a tuberculose bovina é caracterizada como uma zoonose e, em humanos, sua ocorrência é mais observada entre os profissionais que são responsáveis pelo manejo dos animais infectados ou no trato de seus produtos. A tuberculose se profere de duas formas principais: intestinal e pulmonar [52].

### 2.1 Epidemiologia

A tuberculose bovina, através do seu agente *M. bovis* é transmitida aos mamíferos de forma direta ou indireta, não se diferenciando na forma de transmissão entre o bovino e os demais mamíferos [3].

A principal forma de infecção pelo *M. bovis* ocorre pela respiração, e corresponde cerca de 90% das transmissões. A contaminação pelas vias respiratórias se deve a inalação de aerossóis contaminados com a micobactéria. Além da disseminação pela respiração, o bacilo pode ser eliminado pelo leite, fezes, urina, secreção vaginal e sêmen[17] [52].

Além disso, é importante destacar que o bovino infectado pode transmitir o agente *M. bovis* antes mesmo que apareça qualquer tipo de lesão nos tecidos, ou mesmo qualquer outro sintoma clínico [52].

Vale destacar também que os rebanhos de gado leiteiro possuem risco acentuado de transmissão, pois a criação destes bovinos é mais livre e permite o contato direto entre os animais. Visto que o leite pode conter o *M. bovis*, é notório que os bezerros tenham uma grande possibilidade de contrair tuberculose ao ingerir o leite infectado [52].

Embora a contaminação por *M. bovis* não possua qualquer relação com a estação do ano, clima, sexo e região, a idade avançada dos animais eleva a possibilidade de contágio destes animais. [17] [52].

Pode-se dizer que a transmissão do *M. tuberculosis* é quase inexistente no bovino e a contaminação pelo agente *M. avium* possui poucas chances de infectá-los, e se isto ocorrer muito dificilmente a doença progride [48].

A contaminação humana ocorre pelo consumo de leite contaminado; pelo consumo da carne contaminada ou pela inalação de aerossóis. Esta última é a forma mais comum de transmissão da doença, além de ser a mais difícil de ser controlada, e por isso apenas a erradicação pode proteger o criador, sua família e, principalmente, os profissionais que atuam com o rebanho[52].

Em relação ao consumo de carnes e leites infectados, as chances são diminuídas, principalmente porque antes do consumo da carne os frigoríficos são compelidos a realizarem testes com as carcaças do animal. Além disso, a transmissão pelo leite contaminado pode ser reduzida significativamente através do processo de pasteurização. [52].

Em humanos, a tuberculose causada por *M. bovis* é clinicamente indistinguível da tuberculose causada por *M. tuberculosis* e, portanto, deve ser tratada da mesma forma.

Cabe ressaltar que o *M. tuberculosis* é o bacilo mais relevante na causa da tuberculose humana, mas existe uma proporção desconhecida de humanos com tuberculose que é devida à presença do *M. bovis* [38].

## 2.2 Diagnóstico

Para realização do diagnóstico da tuberculose pode-se recorrer aos métodos clínicos, alérgicos, bacteriológicos, sorológicos e/ou anatomopatológicos [55].

O diagnóstico clínico da tuberculose é um processo muitas vezes complicado devido ao fato dos sinais clínicos da doença aparecerem apenas no estágio final da mesma e porque estes sintomas não caracterizarem apenas a tuberculose bovina, mas também outras doenças [55].

Desta forma há a necessidade da realização de testes de diagnósticos. Assim, o teste que utiliza o método de tuberculinização é o mais comum. É um método de investigação alérgico e, até o momento, a forma mais eficaz de diagnosticar a tuberculose em animais vivos, embora possua algumas limitações [53].

Para a realização deste teste, o veterinário pode escolher entre dois locais no corpo do bovino, já que possuem maior sensibilidade que são as regiões do pescoço ou a prega ano-caudal [11].

O Derivado Protéico Purificado (PPD) é a tuberculina mais utilizada. Em geral, é produzido com *M. bovis*, contém 0,1mg/mL (5.000 DI por dose) dessa proteína, enquanto a tuberculina produzida com *M. avium* contém 0,05mg/mL (2.500 DI por dose), ambas devendo ser aplicadas na dose de 0,1mL [17] [52].

É importante lembrar que este teste só pode ser realizado por um médico veterinário, que deve verificar se a tuberculina está refrigerada, e tomar cuidado para que ela não esteja em ponto de congelamento. Ainda, a tuberculina sempre é injetada por via intradérmica, formando assim, uma urna pápula [53].

O uso indevido da tuberculina, ou seja, o armazenamento errado, o transporte irregular e, até mesmo a aplicação imprópria, pode interferir no resultado do teste [44].

Os testes alérgicos tendem a apresentar um número majorado de resultados falsos negativos em animais de idade igual ou superior a seis semanas [7].

Além disto, é importante acentuar que os animais que realizaram o teste de tuberculina acabam apresentando eficácia reduzida em novos testes se este novo teste foi realizado dentro de um intervalo curto (dessensibilização) [44].

Adicionalmente, deve-se verificar se a vaca está prenha ou pariu recentemente, pois essa prova não deve ser realizada no intervalo de 15 dias antes do parto e até 15 dias após o parto, devido a ocorrência de hipossensibilidade neste período. Caso algum animal tenha sido testado para se obter resultados mais precisos estes, devem ser testados novamente no intervalo de 60 a 90 dias após o parto [7].

Por outro lado, bezerros que ingerem colostro de mães infectadas podem apresentar reações positivas ao teste tuberculínico em até três semanas após o nascimento, embora, muito provavelmente, possam não estar infectados [52].

Ainda o veterinário responsável pela aplicação da tuberculina deve verificar se o bovino recebeu algum tipo de aplicação parenteral de soros, hormônios ou medicamentos, pois estes procedimentos podem prejudicar a resposta à reação da tuberculina [55].

Dessa forma, alguns criadores de bovinos se utilizam de algumas dessas práticas para encobrir um resultado positivo, visto que se a tuberculina for injetada durante o período dessensibilização, provocado por estes agentes químicos, é mais incerto apresentar reações positivas [52].

Existem dois tipos de testes tuberculínicos: o teste intradérmico simples, que é realizado somente com o PPD bovino e o teste intradérmico comparativo, no qual utiliza-se, além do PPD bovino, o PPD aviário simultaneamente [44].

Para a realização da prova intradérmica simples cervical, primeiramente determina-se uma área quadrada da pele, com aproximadamente  $9\text{cm}^2$ , situada cranialmente à espinha da escápula.

Após a demarcação da área na região correta, faz-se uma dobra vertical da pele e afere a sua espessura com o auxílio de um cutímetro. Posteriormente, injeta-se o PPD bovino na parte dorsal da prega da pele. De  $72\text{h} \pm 6\text{h}$  após a inoculação é avaliada a reação do animal ao PPD Bovino [7].

Ao avaliar o resultado, mede-se novamente a espessura da pele no local da injeção e, com este resultado, subtrai-se do valor de espessura inicial, obtendo uma medida

que representa o aumento da espessura da pele (em milímetros).

Após a medição, a área que recebeu a injeção de tuberculina deve ser inspecionada pelo veterinário que verificará a existência de outras alterações, como exsudação, edema, endurecimento e necrose no local de aplicação [7] [17] [52].

A avaliação do Teste Cervical Simples é julgando o aumento da espessura da dobra da pele onde se aplicou o PPD bovino. Pode-se notar que na Tabela (2.1), retirada do PNCEBT [7], uma reação negativa é considerada para um aumento da espessura da pele de até 1,9 mm, já uma reação inconclusiva ocorre quando o aumento de 2,0 a 3,9 mm, sem exsudação e necrose, e uma reação positiva necessita ao menos 4 mm de aumento [7].

Tabela 2.1: Interpretação do Teste Cervical Simples em bovinos [7]

<b>Características da Reação</b>				
$\Delta_{Bov}(mm)$	<b>Sensibilidade</b>	<b>Consistência</b>	<b>Outras Alterações</b>	<b>Interpretação</b>
0 a 1,9	-	-	-	Negativo
2,0 a 3,9	pouca dor	endurecida	delimitada	Inconclusivo
2,0 a 3,9	muita dor	macia	exsudto, necrose	Positivo
$\geq 4$	-	-	-	Positivo

Um dos grandes problemas na interpretação do Teste Cervical Simples é a ocorrência de algumas reações inespecíficas, oriundas da presença de outras micobactérias patogênicas[53]. Essas reações podem, por exemplo, ser resultado da tuberculose humana ou aviária [17].

O PNCEBT recomenda a realização da prova comparativa em casos de rebanhos com muito resultados inconclusivo no teste cervical simples. Para isso, deve-se esperar no mínimo 90 dias após a realização do teste cervical simples[7].

Para a realização do teste cervical comparativo, o veterinário deve injetar simultaneamente, mas em locais separados da pele, a PPD bovina e a PPD aviária e avaliar as reações.

Recomenda-se a inoculação da PPD aviária em um ponto da pele situado cranialmente à espinha da escápula e a PPD bovina, cerca de 15 a 20 cm da primeira [7] [17] [53]. É importante destacar que as PPD bovino e aviária devem ser injetadas de um mesmo lado em todos os animais do estabelecimento de criação.

A avaliação é realizada de forma análoga ao teste cervical simples, após 72h  $\pm$  6h da inoculação, afere-se o aumento da espessura da dobra da pele onde aplicou-se PPD bovina e PPD aviária conforme Tabela (2.2) extraída do PNCEBT [7].

Tabela 2.2: Interpretação do teste cervical comparativo em bovinos [7]

	$\Delta B_{ov} - \Delta A_{v}$ (mm)	<b>Interpretação</b>
$\Delta B_{ov} < 2$	-	Negativo
$\Delta B_{ov} < \Delta A_{v}$	$\Delta B_{ov} < 0$	Negativo
$\Delta B_{ov} \geq \Delta A_{v}$	0 a 1,9	Negativo
$\Delta B_{ov} > \Delta A_{v}$	2,0 a 3,9	Inconclusivo
$\Delta B_{ov} > \Delta A_{v}$	$\geq 4$	Positivo

Os exames histopatológicos são realizados utilizando tecidos que são obtidos cirurgicamente ou na necropsia. Este exame possui resultados de diagnósticos positivos se for encontrado a presença de microrganismos do tipo Bacilos Álcool-Ácido Resistentes (BAAR) no interior dos tubérculos[33].

Apesar disto, ainda existe a necessidade do isolamento do microorganismo para que seja estabelecido seu tipo [33]. Assim, o meio de cultura mais recomendado para isolar *M. bovis* é o *Stonebrink* [30].

Fica evidente que o exame histopatológico certamente nos fornece resultados do diagnóstico mais precisos, porém o diagnóstico positivo de tuberculose bovina só é exato se houver o isolamento bacteriológico do *M. bovis*.

Para conseguir o isolamento, e, assim poder identificar o Bacilo *M. bovis* leva-se muito tempo e, em locais onde a prevalência da doença é alta, a necessidade do isolamento por cultura torna-se não viável, e portanto, o diagnóstico dominante seja apenas as avaliações das lesões macroscópicas encontradas durante o exame *post mortem* [16].

## 2.3 Diagnósticos diferenciais

Os principais diagnósticos diferenciais são: doença do trato respiratório superior, linfossarcoma, raiva, edema, paralisia, tumores da laringe, leucose bovina, linfadenopatia, outras causas de mastite e pneumonias micóticas [52] [55].

## 2.4 Prognóstico

O prognóstico para a tuberculose bovina nunca é bom, porque em geral o bovino é diagnosticado tardiamente e o tratamento é proibido por lei. [17].

Portanto, os animais que foram diagnosticados positivos devem ser mantidos isolados e longe do rebanho e só podem manejá-los funcionários com os devidos equipamentos de proteção. Eles devem permanecer no isolamento por um prazo máximo de trinta dias após o diagnóstico. Ao fim deste prazo, estes animais devem

ser sacrificados, em estabelecimento sob serviço de inspeção oficial, indicado pelo serviço de defesa oficial federal e/ou estadual [7].

Na impossibilidade de sacrifício em estabelecimento sob serviço de inspeção oficial, indicado pelo serviço de defesa federal e/ou estadual, os animais serão sacrificados no estabelecimento de criação, sob fiscalização direta da unidade local do serviço de defesa oficial, respeitando procedimentos estabelecidos pelo Departamento de Defesa Animal [7].

Em relação ao sacrifício este deve respeitar alguns critérios preconizados pelo MAPA: deve-se utilizar métodos que causem uma morte rápida e sem espalhar sangue; o abate deve ser feitas preferencialmente no local onde o animal será enterrado, o que evitaria possível contaminação de outros animais ou das pessoas que iriam participar do transporte, e deve-se fazer a cova em terreno seco e estável, distante de poços e cursos de água, nascentes e bebedouros para evitar contaminação. A carcaça deve ser recoberta por um estrato de terra de aproximadamente 2 metros [7].

Cabe lembrar que não é aconselhável realizar necropsia, pois a tuberculose é uma zoonose, mas se houver necessidade, deve-se usar equipamentos de proteção e descontaminá-los fervendo por 30 minutos ou mergulhando os objetos em desinfetantes químicos [7].

## 2.5 Controle e Profilaxia

O controle da tuberculose bovina no Brasil fundamenta-se na periodicidade dos testes preconizados pelo PNCEBT e do sacrifício dos animais que resultarem positivo.

Em áreas de produção de leite existe uma recomendação para que este período seja anual [9] [53]. Em regiões de produção de gado de corte, o período pode ser majorado, visto que existe a possibilidade de identificação dos rebanhos infectados através dos exames histopatológicos das lesões observadas nos locais de abate [17] [53].

No Brasil é realizada a certificação de estabelecimento de criação livre de tuberculose. Este certificado é emitido pela Delegacia Federal da Agricultura, e consta do PNCEBT (Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose Animal), do Ministério da Agricultura, Pecuária e Abastecimento, condicionado à obtenção de três testes tuberculínicos consecutivos e com resultados negativos de cada bovino do rebanho, realizados num intervalo de 90 a 120 dias entre o primeiro e o segundo testes, e de 180 a 240 dias entre o segundo e o terceiro testes. O certificado de estabelecimento de criação livre de tuberculose tem validade de 12 meses [7].

## 2.6 Impacto Econômico da TB bovina

Apesar do crescimento em produção e exportação de produtos agropecuários nos últimos anos, o Brasil vem enfrentando concorrência internacional provocada pelas exigências sanitárias cada vez mais rigorosas que são impostas pelos principais países importadores aos produtos derivados de origem animal oriundos de países sem um consolidado programa de controle da Tuberculose bovina. Este fato limita os Países sem um programa avançado de TB bovina a vender seus produtos por um preço menor a mercados sem possuir tantas barreiras sanitárias[34].

O PNCEBT pretende estimular os setores do agronegócio valendo-se de medidas de adesão voluntária, relativas à certificação das propriedades e que poderiam ser utilizadas como dispositivo pelos criadores para agregar valor aos seus produtos[7].

Um dos principais impedimentos para a adesão em massa pelos criadores em relação as medidas sanitárias propostas pelo PNCEBT é porque o próprio criador tem que se responsabilizar com os custos dos testes, bem como, com a perda do animal em caso de algum destes tiver diagnóstico positivo[7].

Além disto, as principais perdas econômicas desta enfermidade são decorrentes da redução da produção de leite, da redução do ganho de peso e da infertilidade, além é claro da destruição de bovinos que resultem positivos nos testes preconizados pelo PNCEBT[9] [57].

Entretanto, a conduta dos criadores pode ter alguma mudança significativa, em relação à adesão das medidas voluntárias do PNCEBT caso sejam criados subsídios e outros estímulos proveniente do setor privado e, principalmente, do setor público [46].



# Capítulo 3

## Métodos de classificação Supervisionada

As técnicas de classificação podem ser divididas em supervisionada ou não supervisionada. Na classificação supervisionada se conhece, a priori, a classe que gerou cada padrão. Assim, o que se busca é obter funções que são treinadas para classificar de forma correta novos casos. Na classificação não supervisionada os padrões de treino não se encontram classificados, assim, os algoritmos têm por objetivo encontrar uma estrutura matemática nos dados que permita dividi-los em grupos e que minimizem alguma medida de dissimilaridade. Este método se divide em algoritmos hierárquicos ou algoritmos por partição [21].

Como visto acima, a classificação supervisionada está interessada na proposição de uma regra que seja capaz de classificar novos caso, assim, deve-se considerar as seguintes definições[21]:

- Um vetor de dados  $p$ -dimensional  $\underline{x} = (x_1, x_2, \dots, x_p)$ , onde as componentes  $x_i = 1, \dots, p$  são as variáveis características de um objeto; o espaço vetorial de todos os valores possíveis de  $x$  é designado por  $\mathbf{X}$
- Padrões que também podem ser denominados de populações, de classes ou de grupos e, são representados pelos conjuntos  $\Pi_1, \dots, \Pi_g, \geq 2$ .
- Amostra de treinamento, que é uma amostra de padrões das  $g$  classes  $T = \{x_j, y_j\}_{j=1}^n$ , utilizada para construir o classificador
- Amostra de teste, com as mesmas características da amostra treinamento, porém é utilizada para avaliar o desempenho do classificador e, portanto, não são utilizadas na estimação dos parâmetros do classificador.
- Por fim, regra de decisão que é uma função  $F(x) : \mathbf{R}^p \rightarrow \mathbf{C}, \mathbf{C} = \{1, \dots, g\}$ , utilizada para a classificação de uma observação  $p$ -dimensional em uma das  $g$  classes.

Em resumo, o processo de classificação consiste na proposição da regra  $\mathbf{F}$  a partir da amostra  $\mathbf{T}$ , de tal forma que para cada  $x \in \mathbf{X}$ ,  $\mathbf{F}(x)$  designa uma classe em  $\mathbf{C}$ . [60].

## 3.1 Métodos Estatísticos de Classificação Supervisionada

Um subconjunto de técnicas supervisionadas é composto por métodos estatísticos [60]. Estes métodos possuem bons resultados quando seus pressupostos são aceitos; os principais modelos estatísticos de classificação com aprendizagem supervisionada são os modelos discriminante linear de Fisher e discriminante quadrático [60].

### 3.1.1 Análise Discriminante

De acordo com JOHNSON & WICHERN, “*análise discriminante é uma técnica multivariada interessada com a separação de uma coleção de objetos (observações) distintos e que aloca novos objetos em grupos previamente definidos*”[32].

Pode-se dizer que a análise discriminante, diferente de outras técnicas multivariadas, não pode ser caracterizada como uma técnica exploratória, pois determina regras que podem ser utilizadas para classificação de novos objetos, possuindo assim um caráter inferencial.

Na análise discriminante, quando as populações são distintas, a acurácia é elevada, porém, quanto mais as populações (classes ou grupos) se misturam, mais complexo passa a ser o procedimento de alocação correta destes objetos às populações.

Além disso, é essencial que para as funções discriminantes sejam válidas deve-se atender os pressupostos de normalidade das variáveis e homogeneidade das matrizes de variância (discriminante linear). A seguir serão apresentados tais pressupostos.

#### Normalidade

A suposição de normalidade multivariada é pertinente à maioria das técnicas multivariadas, porém não existe um teste de hipótese que a verifique.

Assim, O que ocorre na prática, é aceitar a hipótese que se todas as distribuições marginais possuem aderência à curva normal, então o vetor composto destas variáveis também segue uma distribuição normal multivariado[2].

Deste modo, existem alguns testes estatísticos para a verificação e adequação à curva normal univariada; dentre todos estes testes o mais utilizado é *Kolmogorov-Smirnov*.

## Teste de *Kolmogorov-Smirnov* (KS)

Segundo Conover “*este teste verifica o grau de concordância entre a distribuição de um conjunto de valores amostrais e a determinada distribuição teórica que se quer testar*”[14].

Na verdade a essência deste teste é avaliar se os valores amostrais possam ter vindo de uma população com específica distribuição teórica; neste caso a distribuição normal. Para tanto, compara-se a frequência acumulada observada com os quantis com a que se esperaria da distribuição teórica. Assim a estatística F determina o ponto em que as distribuições, teórica e observada, apresentam maiores divergências, indicando se a distribuição amostral nessa diferença máxima pode ser atribuída ao acaso.

Admitindo que  $F(x_0) = \int_{-\infty}^{x_0} f(x) dx = P(X \leq x_0)$  sendo  $f(x)$  a função de densidade da distribuição normal, as hipóteses serão:

$$H_0 : F = F_0 \quad H_1 : F \neq F_0 \text{ para ao menos um valor de } X.$$

Rejeita-se  $H_0$  se a estatística do Teste T, for maior ou igual a d (os valores de ”d” são tabelados). A estatística do teste é:  $T = \sup |F(X) - S(X)|$ . Pode ainda a hipótese alternativa ser  $H_1 : F < F_0$  ou  $H_1 : F > F_0$ [14].

## Homogeneidade

A suposição da igualdade de matrizes de variâncias é extremamente importante para modelo discriminante linear, pois a desigualdade das matrizes é uma séria violação com comprometimentos severos aos resultados deste modelo. Assim, pode-se construir as hipóteses:

$$H_0 = \Sigma_1 = \Sigma_2 = \dots = \Sigma_g \quad H_1 = \text{pelo menos um } \Sigma_i \text{ é desigual, para } i = 1, 2, \dots, g$$

O teste estatístico para a verificação da homogeneidade é realizado por meio do cálculo do coeficiente M, de acordo com a expressão 3.2[45].

$$M = (n - g) \ln |S| - \sum_{i=1}^g (n_i - 1) \ln |S_i| \quad (3.1)$$

$$M = (n - g) \ln |S| - \sum_{i=1}^g (n_i - 1) \ln |S_i| \quad (3.2)$$

Sendo:

$n_i$  = número de unidades experimentais da i-ésima população  $\frac{\sum_{i=1}^g (n_i - 1) S_i}{n - g}$  matriz de covariância amostral conjunta (*pool*) Ainda, se o fator de escala  $C^{-1}$  na equação 3.3 é introduzido, a quantidade  $MC^{-1}$  é, aproximadamente, uma distribuição Qui-Quadrado com  $\frac{1}{2}(g - 1)p(p + 1)$  graus de liberdade[5].

$$C^{-1} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \left( \sum_{i=1}^g \frac{1}{(n_i - 1)} - \frac{1}{n-g} \right) \quad (3.3)$$

Assim, a hipótese nula é rejeitada quando o  $p$ -valor  $< \alpha$ , onde  $\alpha$  é o nível de significância do teste.

### 3.1.2 Função Discriminante Linear de Fisher

Segundo Johnson e Wichern, “a função discriminante linear consiste, basicamente, em separar duas classes de objetos ou fixar um novo objeto em uma das duas classes”[32].

Ao estruturar o método discriminante linear, este tinha como objetivo classificar padrões através de transformações de observações multivariadas em univariadas, tal que as populações  $\Pi_1, \Pi_2$  fossem separadas em relação às médias das duas populações tanto quanto possível.

Sendo  $\underline{\mu}_1$  a média das observações da população  $\Pi_1$  e  $\underline{\mu}_2$  da população  $\Pi_2$  e, considerando a matriz de covariâncias,  $\Sigma$ , comum para ambas as populações, o que na verdade o modelo discriminante linear propõe é uma combinação linear  $\underline{C}$  que maximizasse a razão entre a “soma de quadrados das distâncias entre as médias das populações e  $\underline{Y}$ ”, bem como, a “variância de  $\underline{Y}$ ” de acordo com 3.4.

$$\underline{C}' \underline{X} = Var(\underline{C}' \underline{X}) = \frac{\left[ \underline{C}' \left( \underline{\mu}_1 - \underline{\mu}_2 \right) \right]^2}{\underline{C}' \Sigma \underline{C}} \quad (3.4)$$

Onde  $\underline{C}'$  é o vetor transposto da combinação linear  $\underline{C}$ .

A razão expressa em 3.4 é maximizada por  $\underline{C} = \Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$ , tendo-se, então, a igualdade 3.5, que é conhecida como função discriminante linear populacional (FDL).

$$FDL(\underline{Y}) = \underline{C}' \underline{X} = \left[ \underline{\mu}_1 - \underline{\mu}_2 \right]' \Sigma^{-1} \underline{X} \quad (3.5)$$

Sabendo-se que muito raramente, os parâmetros populacionais  $\underline{\mu}_1, \underline{\mu}_2$  e  $\Sigma$  não são conhecidos, utilizam-se seus estimadores.

Desse modo, os estimadores de  $\underline{\mu}_1$  e  $\underline{\mu}_2$  são, respectivamente, os vetores de médias amostrais  $\bar{\underline{y}}_1$  e  $\bar{\underline{y}}_2$  e o estimador da matriz de covariâncias conjunta estimada é dada por 3.6, respectivamente.

$$S_{pool} = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2} \quad (3.6)$$

Dessa forma, a função discriminante linear amostral (fdl) fica determinada por 3.7.

$$fdl(\underline{y}) = \left( \bar{y}_1 - \bar{y}_2 \right)' S_{pool}^{-1} \underline{y} \quad (3.7)$$

A estimativa do ponto médio amostral é dada pela expressão 3.8.

$$\hat{m} = \frac{1}{2} \left[ \left( \bar{y}_1 - \bar{y}_2 \right)' S_{pool}^{-1} \left( \bar{y}_1 + \bar{y}_2 \right) \right] \quad (3.8)$$

Ao fim do processo, determina-se a seguinte regra de classificação: Alocar  $\underline{y}$  a  $\Pi_1$  se  $\left( \bar{y}_1 - \bar{y}_2 \right)' S_{pool}^{-1} \underline{y} \geq \hat{m}$ , caso contrário, alocá-lo a  $\Pi_2$ .

### 3.1.3 Função Discriminante Quadrática

Considerando o caso de duas populações,  $\Pi_1$  e  $\Pi_2$ , com as variáveis discriminantes que possuam distribuição normal, porém a matrizes de covariâncias,  $\Sigma_1 \neq \Sigma_2$ , não são homogêneas então a regra de classificação linear deixa de ser válida e deve-se utilizar somente o modelo discriminante quadrático que é determinada como[42]:

Atribua o objeto à população  $\Pi_1$  se a expressão 3.9 for satisfeita e atribua para  $\Pi_2$ , caso contrário (3.10).

$$R_1 : -\frac{1}{2} \underline{y}' \left( \sum_1^{-1} - \sum_2^{-1} \right) \underline{y} + \left( \mu_1' \sum_1^{-1} - \mu_2' \sum_2^{-1} \right) \underline{y} - k \geq \ln \left[ \left( \frac{c(1|2)}{c(2|1)} \right) \left( \frac{p_2}{p_1} \right) \right] \quad (3.9)$$

$$R_2 : -\frac{1}{2} \underline{y}' \left( \sum_1^{-1} - \sum_2^{-1} \right) \underline{y} + \left( \mu_1' \sum_1^{-1} - \mu_2' \sum_2^{-1} \right) \underline{y} - k < \ln \left[ \left( \frac{c(1|2)}{c(2|1)} \right) \left( \frac{p_2}{p_1} \right) \right] \quad (3.10)$$

$c(1|2)$  = custo de má classificação de um indivíduo de  $\Pi_2$  ser classificado na população  $\Pi_1$   $c(2|1)$  = custo de má classificação de um indivíduo de  $\Pi_1$  ser classificado na população  $\Pi_2$

As equações 3.9 e 3.10 produzem regiões de classificação que são definidas por funções quadráticas de  $\underline{y}$ .

Na prática, é muito complicado possuir os parâmetros populacionais em qualquer estudo, então, a função de classificação passa a ser desenvolvida através da substituição dos parâmetros populacionais  $\mu_1, \mu_2, \sum_1$  e  $\sum_2$  por seus estimadores  $\bar{y}_1, \bar{y}_2, S_1$  e  $S_2$ , respectivamente.

Assim,  $\underline{y}$  é atribuído para  $\Pi_1$ , se a equação 3.11 for satisfeita.

$$-\frac{1}{2}\underline{y}'(S_1^{-1} - S_2^{-1})\underline{y} + \left(\underline{\bar{y}}_1'S_1^{-1} - \underline{\bar{y}}_2'S_2^{-1}\right)\underline{y} - k \geq \ln \left[ \left(\frac{c(1|2)}{c(2|1)}\right) \left(\frac{p_2}{p_1}\right) \right] \quad (3.11)$$

Caso contrário  $\underline{y}$  é atribuído para  $\Pi_2$

Ainda, os procedimentos de discriminante linear e quadrático podem ser generalizados para o caso de  $g$  populações  $\Pi_1, \dots, \Pi_g$ , sendo  $g \geq 2$ , cuja as variáveis discriminantes de cada população  $\Pi_i$  possuam distribuição normal multivariada e, claro, respeitando o pressuposto da igualdade das matrizes de variâncias se for modelo linear, porém as descrições destes métodos não serão apresentadas neste trabalho em razão de que o problema a resolver envolve apenas 2 classes. Ademais, o desenvolvimento destes métodos pode ser encontrado em diversas bibliografias.

## 3.2 Redes Neurais

As redes neurais artificiais são sistemas baseados no funcionamento do cérebro humano, formado por unidades de processamento (neurônios) que têm a capacidade de aproximar funções matemáticas complexas, sendo uma ferramenta potente, também, na classificação de padrões.

De um modo geral, a estrutura das redes neurais são compostas pelos neurônios artificiais dispostos em uma ou mais camadas interligadas por conexões. Na maioria dos modelos neurais, estas conexões estão associadas a algum tipo de peso, os quais possuem a responsabilidade de acumular o conhecimento adquirido durante o processo de treinamento. Estes pesos são também chamados de pesos sinápticos ou, simplesmente, sinapses e são os parâmetros que representam o modelo e servem para ponderar a entrada recebida por cada neurônio na rede[6].

### 3.2.1 Neurônio Biológico

De certo modo o cérebro pode ser visto como um computador, pois ele pode ser comparado com um sistema de processamento de informação altamente complexo e que realiza paralelamente diversas funções. O cérebro se auto organiza fazendo com que seus neurônios realizem certos processamentos como reconhecimento de padrões, percepção de sentidos, controle motor e aprendizagem de idiomas com maior eficácia quando comparados aos melhores computadores existente [47].

O neurônio é uma célula que recebe impulsos elétricos a partir de seus dendritos, estes impulsos atuam no corpo celular, que finalmente são transmitidos a outros neurônios através de seu axônio para novos dendritos [6]. A conexão entre um axônio de um neurônio e um dendrito de outro é denominada sinapse. A sinapse é a

unidade funcional básica envolvendo as membranas plasmáticas de dois neurônios, de modo a formar uma junção pontual e orientada do neurônio pré-sináptico para o pós-sináptico [36].

A Figura 3.1 ilustra um neurônio do sistema nervoso central de um vertebrado, onde são também ilustrados e descritos os componentes do mesmo [6], [31].

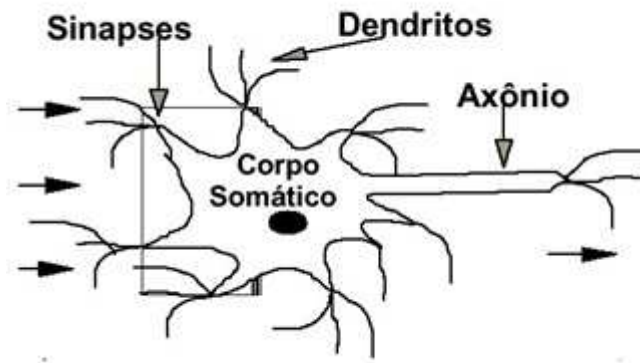


Figura 3.1: Representação de um Neurônio Biológico.

- Corpo Celular do neurônio (ou soma): Onde são realizados os processos metabólicos da célula;
- Dendritos: São as ligações por onde o neurônio recebe informações.
- Axônio: É uma fibra nervosa que serve para conectar o neurônio a outro neurônio;
- Sinapse: Conexão entre um dendrito e um axônio.

### 3.2.2 Neurônio Artificial

Como dito anteriormente, as redes neurais foram fundamentadas na dinâmica do cérebro humano. Os primeiros passos deste novo conceito foram dados por Warren McCulloch e Walter Pitts, considerados por muitos com os criadores do modelo neural na década de 40.

Estes pesquisadores propuseram, matematicamente, o funcionamento de um neurônio biológico que ficou conhecido como modelo de McCulloch & Pitts[41].

Neste modelo, o neurônio possuía  $m$  entradas (dendritos),  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$  e apenas uma saída  $Y$  (axônio). Ainda, com o objetivo de reproduzir o funcionamento das sinapses, cada entrada do neurônio recebeu um peso  $W_1, W_2, \dots, W_m$  associado cujos valores podiam ser positivos (excitatórios) ou negativos (inibitórios).

Deste modo o neurônio era emulado simplesmente somando os valores do produto de suas entradas com seus respectivos pesos,  $X_i W_i \forall i = 1, \dots, m$  e, se a soma dos  $X_i W_i$  for maior ou igual ao seu limiar ( $\Theta$ ) a sua saída é ativada com valor 1.

Resumindo, o neurônio será ativo conforme a expressão 3.12:

$$\sum_{i=1}^m x_i w_i > \Theta \quad (3.12)$$

Onde:

$m$  é o número de entradas do neurônio

$W_i$  é o peso associado à entrada  $X_i$  e

$\Theta$  é o limiar do neurônio, conforme ilustrado na Figura 3.2.

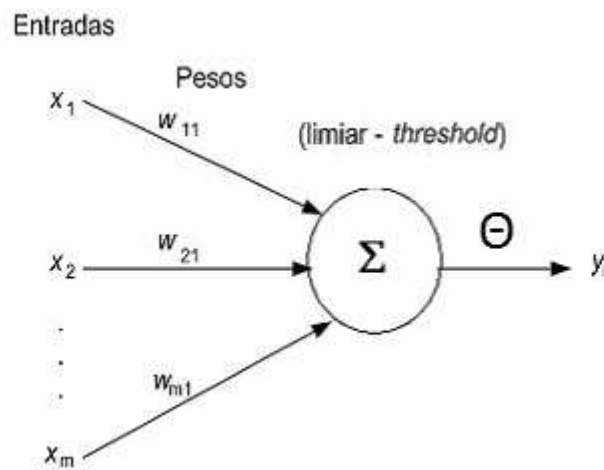


Figura 3.2: Modelo matemático de um neurônio biológico (McCulloch & Pitts [41])

No decorrer dos anos seguintes, vários pesquisadores propuseram variações do modelo do neurônio McCulloch & Pitts que se concentraram, principalmente, em modificações no estado de ativação da saída do neurônio através do uso de funções e na presença de uma componente que corresponde a uma entrada especial. Esta componente é denominada de polarização ou *Bias*. Portanto, o modelo de neurônio artificial que serve de base pra projetos de redes neurais artificiais pode ser representado conforme a Figura (3.3).



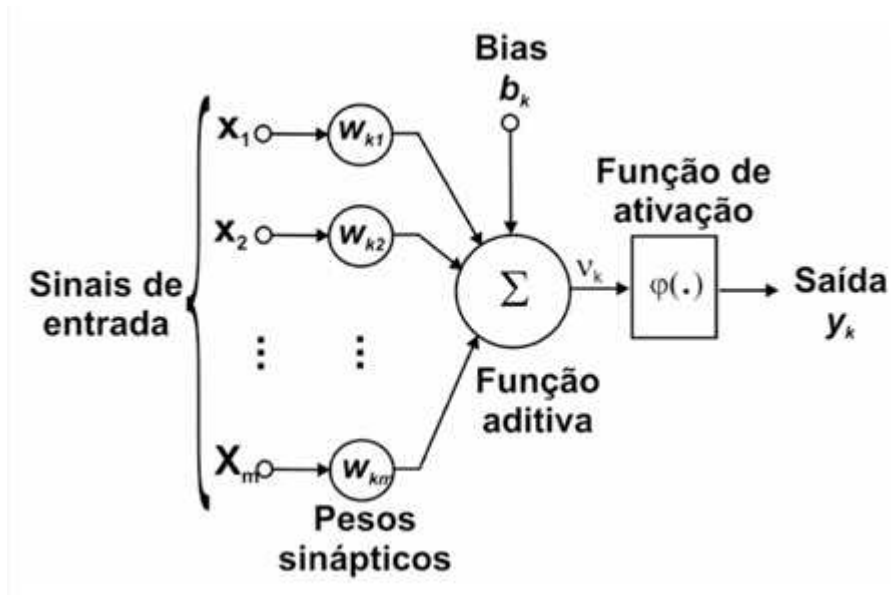


Figura 3.3: Modelo de neurônio base para projetos de redes neurais artificiais

Onde:

- **Pesos Sinápticos** - São considerados os parâmetros do modelo. O papel de uma sinapse  $j$  é conectar um sinal  $x_j$  na entrada a um neurônio  $k$ , já o peso sináptico  $w_{kj}$  tem o propósito de ponderar os sinais de entrada  $x_j$
- **Função aditiva** - Soma dos sinais de entrada  $x_j$  ponderados pelos seus respectivos pesos sinápticos  $w_{kj}$  constituindo um combinação linear como na expressão abaixo:

$$u_k = \sum_{j=1}^m w_{kj}x_j \quad (3.13)$$

- **Função de ativação ( $\varphi(\cdot)$ )**- possui a função de transformar a entrada total recebida por um neurônio em um estado ou sinal de ativação. Em geral, são aplicadas para limitar a amplitude de saída de um neurônio. São usualmente utilizadas quatro tipos básicos de funções de ativação [31]:

1. Função Limiar definida pela equação 3.14.

$$\varphi(v) = \begin{cases} +1, & v > 0 \\ -1, & v \leq 0 \end{cases} \quad (3.14)$$

2. Função Linear definida pela equação 3.15.

$$\varphi(v) = kv \quad (3.15)$$

3. Função logística: É uma função contínua que permite a transição gradual entre os dois estados. É dada pela equação 3.16.

$$\varphi(v) = \frac{1}{1 + e^{-v}} \quad (3.16)$$

4. Função tangente hiperbólica: É uma função sigmóide que varia entre  $(-1, 1)$ . É dada pela expressão 3.17.

$$\varphi(v) = \frac{1 - e^{-v}}{1 + e^{-v}} \quad (3.17)$$

- **Polarização *bias*** - A adição das *bias* no modelo de neurônio artificial tem o propósito de variar a entrada da função de ativação. Assim define-se  $x_0 = 1$  e  $w_{k0}$  como peso sináptico da *bias*. Portanto, define-se  $u_k$  na equação 3.18. A equação 3.19 mostra o sinal de saída do neurônio com a polarização.

$$u_k = \sum_{j=0}^m w_{kj}x_j \quad (3.18)$$

$$\hat{y}_k = \varphi(u_k) \quad (3.19)$$

Onde:

- $x_1, x_2, x_3, \dots, x_m$ , são sinais de entrada;
- $x_0 = 1$ , polarização
- $w_{k0}, w_{k1}, w_{k2}, w_{k3}, \dots, w_{km}$ , são os pesos sinápticos do neurônio  $\mathbf{k}$ ;
- $u_k$  é a saída do combinador linear devido aos sinais de entrada somado a polarização *bias*;
- $\varphi(\cdot)$  é a função de ativação e
- $\hat{y}_k$  é o sinal de saída do neurônio.

### 3.2.3 Características de uma Rede Neural Artificial

Até este momento, foram descritos apenas as características e a evolução de um neurônio artificial. Um neurônio artificial pode ser entendido como uma unidade de processamento. Uma Rede Neural Artificial (RNA) é um modelo que possui diversos destes neurônios artificiais

Existem redes neurais com aprendizado supervisionado, quando é utilizado um padrão resposta esperado para cada vetor de entrada, e o aprendizado não supervisionado, quando não existe um padrão de resposta esperado para cada vetor de entrada.

As RNAs com aprendizagem não supervisionado são capazes de extrair informações não apresentadas de forma explícita e possuem grande capacidade de auto-organização.

Diferentemente, as RNAs com aprendizagem supervisionado utilizam algum algoritmo cujo o objetivo seja minimizar o erro entre o padrão de respostas esperado para cada vetor de entrada e a saída que a rede produz deste respectivo vetor. Em geral, são realizadas inúmeras iterações até que o erro seja minimizado. Ao final do treinamento estas redes armazenam o conhecimento adquirido nos pesos sinápticos, ou seja, possuem capacidade de aprendizagem. Além disto, estas redes treinadas também são capazes de produzir saídas coerentes com o padrão de resposta para vetores de entrada que não participaram do processo de treinamento (generalização).

As redes neurais também possuem boa velocidade de processamento pois, mesmo que exista um grande número de unidades de processamento (neurônios) eles podem operar de forma paralela.

Ainda, uma das características mais significativas das redes neurais é a capacidade de aproximar qualquer função contínua não linear, independente do número de variáveis envolvidas e de possuir bons desempenhos como classificador de padrões.

Ainda, existem diversas arquiteturas de redes neurais artificiais e, o fator fundamental na definição da arquitetura ou da topologia depende do problema a ser tratado. Deste modo, deve-se levar em consideração a quantidades de camadas, o números de neurônios em cada camada e que tipo de conexão será empregada entre os neurônios [6].

Em relação ao número de camadas existem dois tipos de rede:

- **Redes com camada única** - Forma mais simples de uma rede. Possui uma camada de entrada que se projeta para a camada de saída.
- **Redes com múltiplas camadas** - São caracterizadas pela presença de uma ou mais camadas ocultas, tais camadas são postas entre os sinais de entrada e sinais de saída e servem para extrair informações das amostras.

Outra característica das arquiteturas de uma RNA refere-se aos tipos de conexões entre os neurônios:

- **Acíclica ou *feedforward*** - Estruturam as redes de forma que a saída do neurônio em uma determinada camada não possuam conexões com neurônios pertencentes a camadas anteriores nem entre si

- **Cíclicas (*Feedback*)** - Estruturam as redes de forma que a saída do neurônio em uma determinada camada pode ter entradas com neurônios de camadas anteriores (realimentação).

Por fim, a quantidade de neurônios nas camadas ocultas, em geral, é definida empiricamente e sua escolha influencia diretamente na capacidade de aprendizagem e generalização da rede. Uma vez que quantidade baixa de neurônios influencia no tempo de treinamento, uma quantidade superior ao necessário de neurônios pode reduzir o poder de generalização da rede, pois uma quantidade de neurônios em excesso aumenta a possibilidade de memorização dos dados (*overfitting*), principalmente quando há conjunto de dados com poucos elementos. Assim, *overfitting* pode ser entendido como a perda da capacidade de responder a um novo conjunto de dados, apesar da performance do treinamento melhorar.

Para ilustrar uma RNA, a Figura 3.4 apresenta um tipo de estrutura simplificada com conexão do tipo *feedforward* e com apenas uma camada escondida e outra de saída. Este tipo de RNA é conhecida por Multicamadas Perceptrons (MLP) que será descrita na próxima seção.

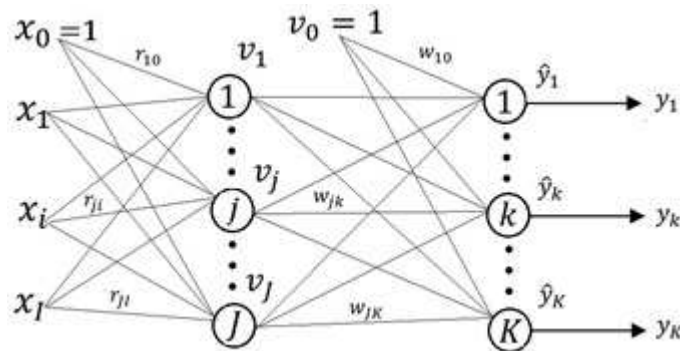


Figura 3.4: RNA de estrutura simplificada *feedforward* e com uma camada escondida e outra de saída.

### 3.2.4 Redes Multicamadas *Perceptron* (MLP)

Uma RNA do tipo *Perceptron* Multicamadas (*Multilayer Perceptron*) é constituída por uma camada de entrada da rede (*input layer*) formada por um conjunto de sinais de entrada e pela componente de polarização *Bias*; uma ou mais camadas intermediárias (*hidden layers*) que, de uma forma geral, usam funções de ativação logística ou tangente hiperbólica e uma camada de saída (*output layer*) cuja determinação da dimensionalidade depende da resposta esperada.

Essas camadas são estruturadas com conexões acíclicas, que, conseqüentemente, propaga o sinal de entrada através da rede, camada a camada, em um sentido progressivo da entrada até a saída e, à exceção da camada de entrada, todas as

outras camadas são compostas por neurônios e, portanto, apresentam capacidade de processamento.

As redes do tipo MLP são redes com aprendizado supervisionado e a presença das não-linearidades distribuídas pelas funções de ativação utilizadas nos neurônios das camadas ocultas, bem como, a alta conectividade presente em redes deste tipo tornam difícil a análise teórica dos parâmetros da rede (pesos sinápticos) [31].

A definição da quantidade de camadas escondidas e da quantidade de neurônios que cada uma delas possui são os aspectos que determinam a complexidade do modelo de RNA e, não há regras simples para determiná-las, sendo muitas vezes determinados de forma experimental. Já os pesos sinápticos são inicializados de forma aleatória utilizando uma distribuição uniforme e são atualizados através de um processo de treinamento que utiliza algum algoritmo de aprendizagem supervisionado[31].

Desta forma, o objetivo do processo de treinamento é obter um conjunto de saídas ( $\hat{y}_p$ ) com base nos sinais de entradas ( $x_p$ ) que ajustam os pesos sinápticos para que o conjunto de saídas convirjam para a um conjunto de valores esperado ( $y_p$ ). Estes conjuntos estão representados pelos seguintes vetores:

$$\begin{aligned} \underline{x}_p &= (x_1, \dots, x_i, \dots, x_I)'_p \\ \hat{\underline{y}}_p &= (\hat{y}_1, \dots, \hat{y}_k, \dots, \hat{y}_K)'_p \\ \underline{y}_p &= (y_1, \dots, y_k, \dots, y_K)'_p \end{aligned} \tag{3.20}$$

Onde  $p$  representa o índice do par entrada-saída.

Porém, antes da realização desse processo de treinamento se faz necessário dividir os dados coletados pois assim, será possível avaliar as capacidades de aprendizagem e generalização características inerentes ao aprendizado supervisionado, portanto o conjunto de dados é dividido em 3 conjuntos:

- **Conjunto de treino:** Possui em geral, cerca de 70% a 80% dos dados coletados. É o conjunto utilizado no treinamento da rede (atualização dos pesos sinápticos).
- **Conjunto de teste :** Possui cerca de 20% a 30% dos dados coletados. É utilizado para verificar a performance preditiva da rede, portanto este conjunto não participa do treinamento.
- **Conjunto de validação:** É um subconjunto do conjunto treino. Este conjunto não participa da atualizações dos pesos sinápticos, porém a performance

deste é avaliada em cada época do treinamento e, caso não melhore em até um determinado número de épocas, o treinamento é interrompido. Assim, o conjunto validação é importante porque verifica a eficácia do aprendizado da rede quanto à capacidade de generalização, resguardando os parâmetros da rede da possibilidade de *overfitting*.

Para avaliar a performance do treinamento realizado pela rede propõem-se a utilização de uma função denominada de função objetivo ( $F$ ) e representa uma medida do erro entre a saída esperada e a saída da rede. Ainda, cabe denominar que uma época é a apresentação de todos os pares (entrada e saída) do conjunto de treino no processo de aprendizado.

De forma geral, pode-se caracterizar o processo de treinamento supervisionado com os seguintes passos:

- **Processamento dos dados:** Os dados de entrada devem ser padronizados com média zero e desvio padrão unitário e os dados de saída devem estar dentro dos limites da função de ativação escolhida para a camada de saída.
- **Definição da rede:** Escolher o número de camadas, quantidade de neurônios presente em cada camada e tipo de função de ativação a ser utilizada.
- **Configuração de parâmetros:** Esta etapa depende do algoritmo de aprendizado escolhido, em geral, define-se os pesos sinápticos iniciais de forma aleatória, taxa de aprendizagem e taxa de momento que influenciarão no tempo de processamento.
- **Escolha do critério de parada:** Número máximo de épocas, um erro mínimo desejado ou número de épocas sem que haja melhora na performance do conjunto validação.
- **Execução do algoritmo de aprendizagem:** O algoritmo *backpropagation* é um dos mais comuns, mas existem outros algoritmos de aprendizagem, como *Resilient Propagation* e o *Levenberg- Marquardt*, por exemplo.

As redes neurais são utilizadas em problemas de modelagem não-linear. Portanto, a função objetivo,  $F$ , expressa uma relação não-linear com os pesos sinápticos da rede.

Em consequência disto, a superfície de erro pode não ser convexa e, portanto, a obtenção da melhor representatividade dos pesos sinápticos recai em um problema de otimização não-linear.

De um modo geral não existe um procedimento para otimização não-linear, porém, um dos métodos mais utilizados na resolução deste problema é o método

do gradiente descendente que realiza no plano da função objetivo  $F$  uma busca por um mínimo. As componentes do gradiente da função objetivo são dadas por:

$$\nabla_{ml}F = \frac{\partial F}{\partial w_{ml}} \quad (3.21)$$

Onde  $F$  é a função objetivo e  $w_{ml}$  o peso sináptico da conexão da  $m$ -ésima camada com a  $n$ -ésima camada.

A orientação do vetor gradiente, no ponto em que é efetuado, é o sentido de máximo crescimento da função sobre a qual o gradiente é calculado. Logo, para determinar o mínimo, deve-se seguir o sentido contrário ao do gradiente. Desta forma, o gradiente da função objetivo é utilizado na atualização dos pesos sinápticos e a regra de aprendizado das sinapses utilizando o método gradiente descendente é dada por:

$$w_{ml}(n+1) = w_{ml}(n) + \Delta w_{ml}(n) = w_{ml}(n) - \mu \frac{\partial F(n)}{\partial w_{ml}} \quad (3.22)$$

Onde  $n$  representa o número da época (iteração) na qual ocorrerá a atualização dos pesos sinápticos e  $\mu$  é a taxa de aprendizado que representa a progressão em direção de um mínimo da função objetivo.

A taxa de aprendizado  $\mu$  é uma constante e em modulo não deve ser muito grande, pois o gradiente descendente requer que sejam tomados passos infinitesimais. Contudo, quanto maior for a taxa de aprendizado, maior será a mudança nos pesos sinápticos, aumentando a velocidade do aprendizado, mas uma taxa com valor elevado pode levar a uma oscilação da rede na superfície da função objetivo.

Em vista disso, o ideal seria utilizar a maior taxa de aprendizado possível que não acarretasse a uma oscilação na superfície da função objetivo, resultando em um aprendizado eficiente mais rápido.

### 3.2.5 Algoritmo *backpropagation*

Redescoberto e divulgado em 1986 por Rumelhart et al., o método *backpropagation* é um procedimento de aprendizagem para redes *feedforward* de múltiplas camadas perceptrons (MLP). A essência do aprendizado *backpropagation* é construir um mapeamento entre entradas e saídas, atualizando os pesos sinápticos. Portanto, espera-se que a rede se torne bem treinada de modo que ela aprenda e possa generalizar [31].

O *backpropagation* (BP) é um algoritmo de otimização de primeira ordem que utiliza o método do gradiente descendente (Equação 3.22) para a correção do valor dos pesos sinápticos da rede.

O algoritmo *backpropagation* baseia-se em uma heurística do aprendizado por

correção de erro, e seu treino consiste em três etapas. A primeira, conhecida por *feedforward*, consiste no cálculo das saídas da rede original por um sinal de entrada. A seguinte etapa compara o sinal de saída resultante do cálculo anterior com o padrão de resposta e retropropaga o erro associado para as camadas anteriores; isto é, no sentido inverso. Na última etapa os pesos são ajustados de acordo com o erro propagado na etapa anterior. Esta sequência é então repetida durante vários ciclos (épocas) de pares de entrada e saída até que um dos critérios de parada seja alcançado.

A função que avalia a performance do treinamento *backpropagation* é erro quadrático médio entre o conjunto saída gerada pela rede e a saída esperada que pode ser escrita como:

$$F_o = \frac{1}{P} \sum_{p=1}^P \sum_{k=1}^K (\varepsilon_{kp})^2 \quad (3.23)$$

$$\varepsilon_{kp} = y_{kp} - \hat{y}_{kp} = y_{kp} - \varphi(u_{kp}) \quad (3.24)$$

Onde  $K$  representa o número de neurônios da camada de saída da rede e  $P$ , o número de pares do conjunto treino. Se a cada época os pesos sinápticos forem atualizados, então a forma no ajuste dos pesos sinápticos é denominada regra delta ou *online*, pois ocorre sempre após apresentação de cada padrão de entrada.

No caso da atualização ocorrer apenas após o processamento de todos os pares entrada-saída, então o método de correção dos pesos sinápticos é denominado de Batelada (*batch*).

O gradiente no treinamento *backpropagation* é aplicado ao erro quadrático. Além disto, uma propriedade importante desse gradiente é que o valor esperado do erro quadrático é igual ao valor esperado do gradiente do erro quadrático, portanto, cada par entrada-saída é tratado isoladamente. Deste modo, o erro quadrático ( $\varepsilon^2$ ) em cada par entrada-saída pode ser descrito:

$$\varepsilon^2 = \sum_{k=1}^K (y_k - \hat{y}_k)^2 \quad (3.25)$$

Assim, o gradiente do erro quadrático em cada par entrada-saída na camada de saída, representada na Figura 3.4, pode ser expresso pela equação 3.26:

$$\nabla_{kj} \varepsilon^2 = \frac{\partial \varepsilon^2}{\partial w_{kj}} \quad (3.26)$$

Já em relação à camada oculta, Figura 3.4, o gradiente do erro quadrático em



cada par entrada-saída é definido pela equação 3.27:

$$\nabla_{ji}\varepsilon^2 = \frac{\partial\varepsilon^2}{\partial r_{ji}} \quad (3.27)$$

Ao aplicar a regra da cadeia à Equação 3.26, o gradiente do erro quadrático em cada par entrada-saída ( $\varepsilon^2$ ) em relação aos pesos sinápticos da camada de saída da rede pode ser expresso:

$$\begin{aligned} \frac{\partial\varepsilon^2}{\partial w_{kj}} &= \frac{\partial}{\partial w_{kj}} \sum_{k=1}^K \varepsilon_k^2 = \sum_{k=1}^K \frac{\partial\varepsilon_k^2}{\partial w_{kj}} = 2 \sum_{k=1}^K \varepsilon_k \frac{\partial\varepsilon_k}{\partial w_{kj}} = 2 \sum_{k=1}^K \varepsilon_k \frac{\partial(y_k - \hat{y}_k)}{\partial w_{kj}} \\ &= -2 \sum_{k=1}^K \varepsilon_k \frac{\partial\hat{y}_k}{\partial w_{kj}} = -2 \sum_{k=1}^K \varepsilon_k \frac{\partial\hat{y}_k}{\partial u_k} \frac{\partial u_k}{\partial w_{kj}} \\ &= -2 \sum_{k=1}^K \varepsilon_k \frac{\partial\hat{y}_k}{\partial u_k} \frac{\partial u_k}{\partial w_{kj}} = -2 \sum_{k=1}^K \varepsilon_k q_{kj} u_k = -2v_j \delta_k \end{aligned} \quad (3.28)$$

Onde:

$$\frac{\partial\hat{y}_k}{\partial u_k} = q_{kj}, \quad u_k = w_{kj}v_j, \quad \frac{\partial u_k}{\partial w_{kj}} = v_j \quad e \quad \delta_k = \sum_{k=1}^K \varepsilon_k q_{kj}$$

O parâmetro  $\delta_k$  representa um sinal de erro retropropagado da saída até a extremidade da sinapse. A atualização do peso sináptico na camada de saída pode ser reescrita como:

$$w_{kj}(n+1) = w_{kj}(n) + 2\mu\delta_k v_j \quad (3.29)$$

Portanto, a equação 3.29 é a regra de aprendizado dos pesos sinápticos dos neurônios da camada de saída da rede neural representada pela Figura 3.4

Analogamente, infere-se a equação de aprendizado para os neurônios da camada oculta da rede da Figura 3.4. Assim, a regra referente aos ajustes dos pesos sinápticos da camada oculta é dado por:

$$\Delta r_{ji} = -\mu \frac{\partial\varepsilon^2}{\partial r_{ji}} = -\mu \frac{\partial\varepsilon}{\partial t_j} \frac{\partial t_j}{\partial r_{ji}} = 2\mu\delta_j x_i \quad (3.30)$$

Com

$$t_j = \sum_{i=0}^I r_{ji} x_i, \quad 1 \leq j \leq J \quad (3.31)$$

$$v_j = \varphi(t_j) \quad (3.32)$$

A derivada de  $t_j$  em relação as sinapses da camada oculta,  $r_{ji}$ , é igual a  $x_i$  e:

$$\delta_j = -\frac{1}{2} \frac{\partial \varepsilon}{\partial t_j} \quad (3.33)$$

Uma das dificuldades do método gradiente descendente era em relação a influência do termo  $t_j$  com os erros da camada de saída, porém ao se aplicar a regra da cadeia, a Equação 3.33 pode ser reescrita como:

$$\delta_j = -\frac{1}{2} \frac{\partial \varepsilon}{\partial v_j} \frac{\partial v_j}{\partial t_j} \quad (3.34)$$

Cuja a derivada parcial de  $v_j$  em relação a  $t_j$  é representada por  $\varphi'(t_j)$ .

A derivada parcial da função  $\varepsilon$  em relação a  $v_j$  pode ser definida como:

$$\frac{\partial \varepsilon}{\partial v_j} = \frac{\partial}{\partial v_j} \sum_{k=1}^K [y_k - \varphi(u_k(\underline{v}))] = -2 \sum_{k=1}^K [y_k - \hat{y}_k] \frac{\partial}{\partial v_j} \varphi(u_k(\underline{v})) \quad (3.35)$$

Onde  $\underline{v}$  expressa o vetor contendo as saídas dos neurônios da camada oculta e é dado por:

$$\underline{v} = [v_1, \dots, v_J]' \quad (3.36)$$

O vetor  $\underline{v}$  foi escrito na Equação 3.35 para mostrar a dependência das saídas dos neurônios da camada de saída em relação às saídas da camada oculta. A regra da cadeia na derivada parcial da Equação 3.35 determina que:

$$\frac{\partial}{\partial v_j} \varphi(u_k(\underline{v})) = \frac{\partial}{\partial u_k} \varphi(u_k(\underline{v})) \frac{\partial u_k}{\partial v_j} = \varphi'(u_k) w_{kj} \quad (3.37)$$

Desta forma, a Equação 3.35 pode ser formulada como:

$$\frac{\partial \varepsilon}{\partial v_j} = -2 \sum_{k=1}^K \varepsilon_k \varphi'(u_k) w_{kj} = -2 \sum_{k=1}^K \delta_k w_{kj} \quad (3.38)$$

Conhecidas as Equações 3.34 e 3.38, conclui-se que a expressão singular do erro na camada oculta como sendo:

$$\delta_j = \varphi'(u_j) \sum_{k=1}^K \delta_k w_{kj} \quad (3.39)$$

Dadas as Equações 3.30 e 3.39, se pode expressar a equação de aprendizado das

sinapses da camada oculta da rede neural representada na Figura 3.4 como sendo:

$$r_{ji}(n+1) = r_{ji}(n) + 2\mu x_i \varphi'(u_j) \sum_{k=1}^K \delta_k w_{kj} \quad (3.40)$$

As derivadas das funções de ativação dos neurônios das camadas ocultas e de saída são indispensáveis para o cálculo dos erros,  $\delta$ . Se o neurônio possuir função de ativação tangente hiperbólica (3.17), sua derivada pode ser expressa, como:

$$\varphi'(u) = 1 - \hat{y}^2 \quad \varphi'(t) = 1 - v^2 \quad (3.41)$$

O Algoritmo *backpropagation* teve seu procedimento de cálculo resumido para uma rede MLP com 1 camada oculta e utilizando o método Batelada na atualização dos pesos sinápticos. Neste método, o ajustes das sinapses só ocorre após todos os pares terem sido processados pela rede.

1. Escolher o passo de treinamento  $\mu$ , dentro do intervalo  $(0; 1)$ , o valor de erro desejado  $F_0$  ou estabelecer um número máximo de iterações,  $N$ ;
2. Inicializar os valores das sinapses das camadas ocultas e de saída de forma aleatória seguindo uma distribuição uniforme  $[-0, 2; 0, 2]$ ;
3. Treinar a rede com o próximo par entrada-saída  $(x_p, y_p)$ ;
4. Calcular as saídas das camadas intermediária e de saída da rede;
5. Calcular a soma dos erros quadráticos de acordo com o par de treinamento utilizado e com as sinapses das camadas intermediária e de saída:

$$F_p = \sum_{k=1}^K (y_{pk} - \hat{y}_{pk})^2 \quad (3.42)$$

Caso  $F_0$  tenha sido deliberado testar, se  $F_p < F_0$ . Se verdadeiro, terminar o algoritmo. Caso contrário ir para o item 6.

Os valores das sinapses  $w_{kj}$  e  $r_{ji}$  são constantes durante o treinamento dos  $P$  pares:

6. Calcular os valores das atualizações das sinapses da camada de saída da rede para  $k = 1, \dots, K$  e para  $j = 0, \dots, J$ :

$$\Delta w_{kj} = \sum_{p=1}^P (y_{kp} - \hat{y}_{kp}) \varphi_{\hat{y}_k}(u_{kp}) v_{jp} \quad (3.43)$$

7. Calcular os valores das atualizações das sinapses da camada intermediária da rede para  $j = 1, \dots, J$  e para  $i = 0, \dots, I$ :

$$\Delta r_{ji} = \sum_{p=1}^P x_{ip} \varphi'_{hj}(u_{jp}) \sum_k k = 1 \delta_{kp} w_{kjp} \quad (3.44)$$

8. Atualizar os valores das sinapses da camada de saída da rede:

$$w_{kj} = w_{kj} + \frac{2}{P} \mu \Delta w_{kj} \quad (3.45)$$

9. Atualizar os valores das sinapses da camada escondida da rede:

$$r_{ji} = r_{ji} + \frac{2}{P} \mu \Delta r_{ji} \quad (3.46)$$

10. Se  $p < P$  voltar para o item 3. Se  $p = P$  fazer  $n = n + 1$  e se  $n \geq N$  terminar o algoritmo. Caso contrário voltar para o item 3.

A taxa de aprendizado ( $\mu$ ) interfere na velocidade de convergência da rede. Se esta for muito pequena, a quantidade de épocas necessárias para alcançar um modelo satisfatório será grande. Por outro lado, o uso de um valor alto poderá ocasionar oscilações que dificultam a convergência. Assim, para amenizar tal situação, é utilizado o termo *momentum* quando utilizado no processo *online* (Regra Delta), o que quantifica o grau de importância da variação de peso da etapa anterior à atual. Este procedimento pode ser considerado intermediário entre treinamento Regra Delta e treinamento por batelada.

Assim, os pesos sinápticos da camada de saída e da camada oculta passam a ser atualizados através das equações 3.47 e 3.48 respectivamente:

$$w_{kj}(n+1) = w_{kj}(n) + 2\mu \delta_k v_j + \alpha(w_{kj}(n) - w_{kj}(n-1)) \quad (3.47)$$

$$r_{ji}(n+1) = r_{ji}(n) + 2\mu x_i \varphi'(u_j) \sum_{k=1}^K \delta_k w_{kj} + \alpha(r_{kj}(n) - r_{kj}(n-1)) \quad (3.48)$$

Em geral, utiliza-se  $\alpha$  aproximadamente a 0,9.

### 3.2.6 Algoritmo *Resilient Propagation*

O algoritmo *Resilient propagation* (Rprop) foi desenvolvido com o intuito de acelerar o processo de aprendizagem. A ideia básica é ajustar o passo de treinamento  $\mu$

conforme a permanência ou mudança do sinal do gradiente nas sucessivas iterações. Com esta mudança se adaptaria a influência desta magnitude nos ajustes de pesos, visto que a derivada de funções sigmóides são pequena para pontos afastados de zero o que gera gradientes locais muito pequenos ocasionando uma lentidão na minimização da função objetivo [31]. Para estes pontos afastados a polaridade do gradiente é normalmente constante, enquanto que para pontos próximo ao minimante ela varia. Uma variação na polaridade do gradiente indica que o minimante foi ultrapassado, i.e., que acréscimo foi excessivo.

Portanto, o ajuste da taxa de aprendizado ( $\mu$ ) se dá através dos sinais dos gradientes do erro quadrático ( $\varepsilon^2$ ), são adaptados uma única vez a cada época ( $n$ ) de treinamento. Cada peso possui sua própria taxa de variação ( $\mu_{kj}$ ), a qual se altera mediante a época ( $n$ ) da seguinte forma:

$$\mu_{kj}(n) = \begin{cases} a\mu_{kj}(n) & , \quad se \frac{\partial \varepsilon^2(n-1)}{\partial w_{kj}} \cdot \frac{\partial \varepsilon^2(n)}{\partial w_{kj}} \geq 0 \\ b\mu_{kj}(n) & , \quad se \frac{\partial \varepsilon^2(n-1)}{\partial w_{kj}} \cdot \frac{\partial \varepsilon^2(n)}{\partial w_{kj}} < 0 \end{cases} \quad (3.49)$$

Onde tipicamente  $a \approx 1,05$  e  $b \approx 0,95$

Observa-se que a mudança nos passos de treinamento depende apenas do sinal das derivadas parciais, independendo de sua magnitude. Caso a derivada parcial não seja negativa, o peso será incrementado, porém, se a derivada for negativa, o peso será decrementado.

Pode-se, então, deduzir que a regra de aprendizagem dos pesos sinápticos é também baseada na mudança de sinais das derivadas parciais dos pesos sinápticos correspondentes. Assim uma mudança de sinal na derivada parcial indicará que a última atualização foi muito grande e saltou uma região de mínimo local. Se isto ocorrer  $\mu_{kj}$  será decrementado. Por outro lado, se a derivada parcial mantiver o sinal,  $\mu_{kj}$  será ligeiramente acrescido para auxiliar na aceleração da convergência em uma região rasa.

Uma vez que o valor de atualização para cada peso é adaptado, a atualização dos pesos sinápticos segue uma regra muito simples: se a derivada é positiva, o peso sináptico é diminuído pelo seu valor de atualização, se a derivada for negativa, o valor de atualização é acrescido:

$$\Delta_{ij}(n) = \begin{cases} \mu^+ * \Delta_{ij}(n-1) & , \quad se \frac{\partial \varepsilon^2(n-1)}{\partial w_{ij}} * \frac{\partial \varepsilon^2(n)}{\partial w_{ij}} > 0 \\ \mu^- * \Delta_{ij}(n-1) & , \quad se \frac{\partial \varepsilon^2(n-1)}{\partial w_{ij}} * \frac{\partial \varepsilon^2(n)}{\partial w_{ij}} < 0 \\ 0 & , \quad caso \text{ contrário} \end{cases} \quad (3.50)$$

$$w_{ij}(n+1) = w_{ij}(n) + \Delta_{ij}(n) \quad (3.51)$$

Para evitar divergência no treinamento deve-se ainda manter  $\mu$  limitado. Se  $\mu_{kj} > \mu^*$ , então  $\mu_{kj} = \mu^*$ . Onde tipicamente  $\mu^* = 0,2$

Se a derivada parcial muda de sinal, isto é, o passo anterior foi muito grande e a região mínima foi ultrapassada, a atualização do peso pode ser revertida:

$$\Delta w_{ij}(n) = -\Delta w_{ij}(n), \text{ se } \frac{\partial \varepsilon^2(n-1)}{\partial w_{ij}} * \frac{\partial \varepsilon^2(n)}{\partial w_{ij}} < 0 \quad (3.52)$$

Portanto, a lógica do algoritmo Rprop é que se a derivada parcial manteve o sinal, o ajuste foi insuficiente e deve ser ampliado, e caso contrário, se houve alteração do sinal da derivada, o ajuste foi superior e deve ser reduzido; tal lógica induz a um processamento mais ágil[31].

### 3.3 Métodos de Avaliação dos modelos Supervisionados

A acurácia refere-se ao grau em que um modelo de classificação é capaz de determinar o verdadeiro valor do que está sendo medido, ou seja, a acurácia é um índice que mede a taxa de acerto global de um modelo de classificação. A sensibilidade de um modelo de classificação é a capacidade que este possui em detectar os indivíduos verdadeiramente positivos, ao passo que, a especificidade é a capacidade que o modelo de classificação tem de detectar os verdadeiros negativos, isto é, de diagnosticar corretamente os indivíduos sadios.

A Figura 3.5 mostra a matriz de validade de um teste de diagnóstico ou dos resultados de um modelo de classificação quando comparadas ao “padrão ouro”, classificação correta.

Os seguintes índices podem ser calculados a partir dos resultados do quadro demonstrado na Figura 3.5.

$$\begin{aligned} \text{Acurácia} &= \frac{VP+VN}{(VP+VN+FN+FP)} \\ \text{Sensibilidade} &= \frac{VP}{(VP+FN)} \\ \text{Especificidade} &= \frac{VN}{(VN+FP)} \end{aligned} \quad (3.53)$$

Existe, ainda, o índice soma-produto, baseado nos resultados de sensibilidade e especificidade que é calculado de acordo com a equação 3.54.

$$\text{SP} = \sqrt{\frac{S + (FA)}{2}} \times \sqrt{S \times (FA)} \quad (3.54)$$

		Doença ("Padrão Ouro")	
		Positivo	Negativo
Testes de Diagnóstico/ Modelos de Classificação	Positivo	VP (Verdadeiro Positivo)	FN (Falso Negativo)
	Negativo	FP (Falso Positivo)	VN (Verdadeiro Negativo)

Figura 3.5: Validade de um teste de diagnóstico ou dos resultados de um modelo de classificação.

Onde S é a sensibilidade e FA é falso alarme (1-especificidade) do processo de diagnóstico .

### 3.3.1 Curva ROC

A relação entre sensibilidade e especificidade é antagônica e uma maneira de expressá-la é construindo uma curva denominada de curva ROC (*Receiver Operating Characteristic*), que é uma ferramenta poderosa, por exemplo, para medir e especificar problemas no desempenho do diagnóstico médico [24].

A curva ROC, em geral, é uma medida não-paramétrica e um dos seus propósitos está relacionada em avaliar e comparar a potencialidade de discriminação de diversos classificadores.

A construção da curva ROC só é possível após determinar, de forma arbitrária, diversos pontos de cortes e calcular a sensibilidade e especificidade respectivamente destes pontos de corte. Desse modo, a Curva ROC é um gráfico de sensibilidade, representada no eixo das ordenadas versus o índice falso alarme (1 - especificidade), representado no eixo das abscissas.

Um dos critérios de avaliação da curva ROC é em relação a área sob a curva podendo variar de 0,5 a 1. Assim, se um modelo possui sensibilidade e especificidade próximas de 100%, a área sobre a curva se aproxima de 1. Quando a performance de um modelo se torna mais aleatória, a área sob a curva se aproxima-se de 0,5. O cálculo desta área depende da estatística U de Mann-Whitney ou dos resultados do teste de Wilcoxon que são distribuições não-paramétricas utilizadas para casos de

duas amostras independentes.

O gráfico da curva ROC é limitado tanto no eixo das ordenadas quanto no eixo das abscissas no intervalo  $[0, 1]$  uma vez que tanto a sensibilidade quanto a especificidade só assumem valores dentro deste intervalo.

Existe a possibilidade de utilizar uma diagonal na curva ROC, que seria uma referência que indicaria que o modelo não traz nenhuma informação, ou seja, a área sob esta diagonal é 0,5, como dito anteriormente. Modelos que possuem tal característica são modelos com performance muito ruim na discriminação, pois esta se torna randômica [24].

Pelo exposto acima o melhor modelo é o que mais se aproxima da extremidade superior e esquerda do gráfico. Ainda, é muito improvável que um modelo possua sensibilidade e especificidade iguais a 100%. É usualmente aceito que a classificação tem desempenho razoável se a área sob a curva ROC está entre 0,7 a 0,8; bom desempenho quando a área é superior a 0,8 e inferior a 0,9 e excelente quando a área é superior a 0,9 [24].

Um outro propósito da curva ROC é estar na novas regras de decisão, pois cada ponto de corte da curva representa uma nova regra de decisão para o classificador. A sensibilidade, especificidade e, conseqüentemente, a acurácia variam de acordo com ponto de corte.

Assim, a escolha do ponto de corte depende de uma decisão entre aumentar a sensibilidade à custa de redução da especificidade, ou vice-versa. A estratégia geral para isto é a seguinte:

- Se a principal preocupação é reduzir resultados falso-positivo, então o ponto de corte deve objetivar o máximo de especificidade.
- Se a preocupação for reduzir resultado falso-negativo, então o ponto de corte deve objetivar o máximo de sensibilidade.
- Existe também a possibilidade de maximizar o índice SP em vista de aumentar a acurácia do modelo

Por todas estas possibilidades a curva ROC é uma das melhores estratégias de estabelecer o ponto de corte.

### 3.3.2 Validação Cruzada

A tarefa de validação de modelos supervisionados é um processo de determinação do grau de confiança que estes modelos possuem em relação ao conjunto de dados apresentados, principalmente, se estes modelos possuem uma boa capacidade preditiva quando utilizados um conjunto de dados novos.



Dentro deste contexto, uma ferramenta padrão em estatística é conhecida como validação cruzada (*cross-validation*) e fornece um procedimento atrativo, particularmente quando à quantidade de dados rotulados (valor desejado) é precária.

Dentre alguns métodos de validação cruzada se destaca o *multifold-cross-validation* que divide o conjunto de dados disponíveis de  $N$  exemplos em  $r$  subconjuntos,  $r > 1$ . O modelo é treinado na maioria dos subconjuntos, e os índices de desempenho como acurácia sensibilidade e especificidade, por exemplo, são medidos sobre os subconjuntos deixados de fora [31].

Este procedimento é repetido por um total de  $H$  tentativas, cada vez usando subconjuntos diferentes. A performance do modelo é verificada estimando uma média e calculando o erro amostral dos índices de desempenho para os subconjuntos que não participaram do treinamento em cada uma das  $H$  tentativas do experimento. Uma desvantagem deste método é que o modelo tem de ser treinado  $H$  vezes [31].

# Capítulo 4

## Metodologia

Para o presente estudo foram utilizadas informações de 127 bovinos provenientes de uma fazenda em Macaé e da fazenda modelo da UFF. Em todos os animais foram realizados os testes Intradérmicos (Teste Cervical Simples e Teste Cervical Comparativo) e exames confirmatórios *post-mortem* nos tecidos dos animais (histopatológicos) [7].

Deste modo, a variável  $\Delta Bov$  foi obtida aferindo o aumento da espessura da pele do bovino após o período de inoculação do PPD bovino. O Teste Cervical Simples utiliza critérios que avaliam este aumento para classificar o bovino quanto à tuberculose. A Tabela 4.1 descreve estes critérios [7].

Tabela 4.1: Critério de classificação dos resultados do teste cervical simples

$\Delta Bov(\text{mm})$	<b>Interpretação</b>
$\Delta Bov < 2$	Negativo
$2,0 \geq \Delta Bov < 3,9$	Inconclusivo
$\Delta Bov \geq 4$	Positivo

A variável  $\Delta Bov - \Delta Av$  foi obtida efetuando a diferença no aumento das espessuras ocorridas na pele do bovino após o período da inoculação do PPD bovino e PPD aviário aplicados simultaneamente. Assim, o Teste Cervical Comparativo avaliou esta diferença baseado nos critérios descritos na na Tabela 4.2 para classificar quanto à tuberculose bovina[7].

Tabela 4.2: Critério de Classificação dos Resultados do Teste Cervical Comparativo

$\Delta Bov - \Delta Av$ (mm)		Interpretação
$\Delta Bov < 2$	-	Negativo
$\Delta Bov < \Delta Av$	$\Delta Bov < 0$	Negativo
$\Delta Bov \geq \Delta Av$	0 a 1,9	Negativo
$\Delta Bov > \Delta Av$	2,0 a 3,9	Inconclusivo
$\Delta Bov > \Delta Av$	$\geq 4$	Positivo

Ao final da execução dos testes cervicais, os resultados destes foram tabulados e o veterinário responsável do estudo estabeleceu uma variável qualitativa, baseado em testes confirmatórios (histopatológicos).

Essa variável foi denominada status e classificou cada bovino do estudo quanto ao diagnóstico de tuberculose: bovino infectado com tuberculose e bovino não infectado com tuberculose. Portanto, neste estudo, esta variável foi considerada como “padrão ouro” para definição das classes quanto ao diagnóstico de tuberculose.

Deste modo, as variáveis  $\Delta Bov$ ,  $\Delta Bov - \Delta Av$ , os resultados dos testes TCS e TCC e a variável status foram as variáveis utilizadas neste trabalho.

Cabe dizer que foram retirados do estudo 11 bovinos com status não infectado com TB bovina, este procedimento foi realizado para equilibrar as populações de bovinos infectados e bovinos não infectados, procedimento importante, principalmente, para o aprendizado das redes neurais.

Inicialmente, foram avaliados os resultados de diagnósticos quanto à tuberculose dos bovinos deste estudo obtidos pelos testes cervicais que são os preconizados pelo Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose no Brasil (PNCEBT).

Esta avaliação foi realizada através de uma análise cruzada entre cada um dos testes de diagnósticos TCS e TCC com a variável status e obtidos os resultados de taxa de acerto global (acurácia), sensibilidade e especificidade.

Além disso, foi utilizada a curva ROC para cada um das duas variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  que avaliou o poder discriminatório dos testes de diagnóstico TCS e TCC respectivamente, e expressou a relação entre sensibilidade e especificidade destas provas de diagnóstico.

Em seguida, buscou-se um melhor ponto de corte para cada uma das variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  que representasse uma nova regra de decisão de cada um dos testes de diagnóstico.

Estes pontos de corte foram gerados de forma arbitrária e a escolha do melhor destes pôde ser realizado por diversos critérios conforme já explicado.

Dessa forma, definiu-se como critério de escolha pontos de corte que maximizariam a sensibilidade e pontos de corte que maximizariam o índice SP, este segundo

com a finalidade de maximizar a acurácia.

Após a avaliação destas novas regras de decisão pra os testes cervicais, adotou-se uma abordagem multivariada e foram propostos a utilização de métodos estatísticos e de redes neurais na tentativa de obter uma combinação das variáveis  $\Delta Bov$ ,  $\Delta Bov - \Delta Av$  que aumentariam a acurácia do diagnóstico da Tuberculose Bovina e modelos que buscassem aumentar a sensibilidade nos diagnósticos da TB bovina e que evita-se uma forte queda da especificidade.

Assim, optou-se por empregar técnicas de classificação supervisionada visto que a variável status, considerada “padrão ouro”, e que define as classes quanto ao diagnóstico de tuberculose bovina, possui relação de dependência com as variáveis  $\Delta Bov$ ,  $\Delta Bov - \Delta Av$ .

Em geral, nos modelos de classificação supervisionadas há sempre uma única variável dependente que é também transformada em variável categórica. Há, ainda, as variáveis independentes, que podem ser quantitativas ou qualitativas para a construção dos modelos.

Cabe ressaltar que estas técnicas de classificação supervisionadas, são procedimentos que além de distinguirem os bovinos do conjunto de dados quanto ao diagnóstico de tuberculose através de regras e funções de classificação, possuem a capacidade de, conhecendo os resultados das variáveis independentes de um novo bovino, seja possível classificá-lo também.

Alguns estudos sugerem como principais técnicas de classificação supervisionadas a análise discriminante e as redes neurais que têm habilidade de lidar com efeitos não lineares e descontínuos, identificando relações que os métodos estatísticos usuais não contemplam [31].

Antes de prosseguir com o processamento das técnicas de classificação supervisionada estatística houve a necessidade de realizar a etapa de pré-processamento. Assim, os valores da variável Status foram categorizados: N - bovino não infectado com tuberculose e T - bovino infectado com tuberculose.

Nos projetos das redes neurais a variável status foi transformada na variável  $y$  e seus valores em valores numéricos: -1 - bovino não infectado e 1 - bovino infectado, pois o aprendizado das redes neurais utilizou o erro quadrático médio (EQM) entre a estimativa de saída e a saída real para atualizar os pesos sinápticos e os *bias* que são os parâmetros da rede neural.

Ainda, as variáveis  $\Delta Bov$ ,  $\Delta Bov - \Delta Av$  (variáveis contínuas) foram padronizados para média nula e variância unitária (*z-score*).

A Análise Discriminante foi aplicada neste trabalho utilizando a variável status, já transformada em categórica (N- bovino não infectado ; T- bovino infectado) como variável dependente; o  $\Delta Bov$  e o  $\Delta Bov - \Delta Av$  como variáveis independentes.

Para seguir com a avaliação e utilização dos modelos discriminantes foi necessária

a observação dos pressupostos para sua utilização. Assim, os principais pressupostos da análise discriminante que foram verificados: normalidade multivariada e homogeneidade de matrizes de variância.

Para avaliar a normalidade multivariada das variáveis independentes ( $\Delta Bov$ ,  $\Delta Bov - \Delta Av$ ) utilizou-se o teste de Kolmogorov-Smirnov (KS) para uma amostra, visto que não existe um teste de hipóteses para testar aderência à curva normal multivariada e aceita-se que caso as distribuições univariadas possuam aderência à curva normal então a distribuição conjunta também possui aderência à curva normal.

Para a verificação da homogeneidade de variâncias, o teste denominado Box's M testou a hipótese nula de igualdade de matrizes de variância dos grupos (bovino infectado tuberculose e bovino não infectado tuberculose).

Cabe ressaltar que o nível de significância utilizado para cada um dos dois testes foi de 5%.

Após a verificação dos pressupostos, a massa de dados com 116 bovinos foi separada em dois grupos para estimação dos parâmetros dos modelos discriminante linear e quadrático. Assim, o conjunto treino continha 47 bovinos com infecção à tuberculose e 47 sem infecção à tuberculose; o conjunto teste continha 11 bovinos com infecção à tuberculose e 11 bovinos sem infecção à tuberculose.

Esse procedimento de separar em dois grupos a massa de dados foi útil para verificação do poder preditivo dessas funções discriminantes.

As funções discriminantes linear e quadrática foram construídas, porém só seriam consideradas válidas caso fossem aceita a hipótese da normalidade das variáveis independentes, caso contrário, estas funções não seriam significativas.

Ainda, foram construídas as duas funções discriminantes: linear e quadrática, porém de acordo com resultado do teste Box's M, a função discriminante linear poderia perder a significância estatística na determinação dos bovinos quanto ao seu diagnóstico. Apesar disto, optou-se, neste trabalho, por apresentar os resultados dos dois modelos, independente do resultado do teste Box's M, como forma de comparar as duas funções.

Dentre as duas metodologias de classificação supervisionadas apresentadas neste trabalho, a das redes neurais é bem mais flexível que a análise discriminante, pois nenhum pressuposto precisa ser verificado. As redes neurais não pressupõem um modelo ao qual os dados devam ser ajustados, já que o modelo é gerado pelo processo de aprendizagem.

Na modelagem das redes neurais, a massa de dados com 116 bovinos foi separada em 3 grupo para o seu treinamento. O conjunto treino que continha 41 bovinos com status infectado com tuberculose e 41 não infectado com tuberculose; o conjunto teste que continha 11 bovinos com status infectado com tuberculose e 11 bovinos com status não infectado com tuberculose e o conjunto validação com 6 casos de

bovinos com status infectado com tuberculose e 6 bovinos com status não infectado com tuberculose.

O conjunto de treinamento contém os dados que foram mostrados à rede e foi com base neles que a rede foi treinada. O conjunto de validação serviu como uma parada antecipada, visando evitar o problema de *overtraining* e aumentar a capacidade de generalização da rede. O terceiro conjunto, o de teste, serviu para avaliar o desempenho preditivo da rede em dados desconhecidos, sendo, assim, possível avaliar a capacidade de generalização do modelo.

Assim, foram projetadas redes neurais do tipo Perceptron Multicamadas (MLP) com uma camada escondida e realizando o processo de aprendizagem através do algoritmo *resilient propagation* por batelada e com até 2.000 épocas de treinamento. As entradas destas redes foram as variáveis  $\Delta Bov$ ,  $\Delta Bov - \Delta Av$  padronizadas e a saída, a variável status transformada em  $y$  (-1- bovino não infectado com tuberculose; 1- bovino infectado com tuberculose).

Em relação a topologia da rede, pode-se dizer que redes neurais com poucos neurônios na camada escondida são preferidas, visto que elas tendem a possuir um melhor poder de generalização, reduzindo o problema de superestimação (*overfitting*), principalmente quando há restrição estatística no conjunto de dados.

Desta maneira, foram projetadas redes neurais com 1, 2, 3 e 5 neurônios na camada escondida e 1 neurônio na camada de saída. As funções de transferência utilizadas foram a tangente hiperbólica tanto na camada intermediária, pois consolida o processo não linear da rede, quanto na camada de saída, função utilizada para estudos de classificação.

Ainda, para cada topologia de rede foram realizadas 100 inicializações de pesos sinápticos diferentes. Os pesos sinápticos em cada uma das 100 inicializações foram determinados de forma aleatória dentro do intervalo  $(-0, 2; 0, 2)$ .

A determinação do melhor projeto de rede levou em consideração dois critérios: menor erro quadrático médio ( $EQM$ ) e uma boa capacidade de generalização da rede que poderia ser verificada pelo desempenho no conjunto validação.

Além disso, foram projetadas redes neurais com modificações no erro médio quadrático ( $EQM$ ); tais modificações possuem o propósito de penalizar com maior rigor erros de classificação que gerassem falsos negativos, pois assim, os pesos sinápticos seriam atualizados com propósito de aumentar a sensibilidade, que por sua vez reduziria a médio e longo prazo novos casos de bovinos infectados.

Para tanto, foi implementada uma função ponderadora que ao multiplicar o  $EQM$ , incrementaria a penalidade por um fator  $M$  quando o modelo apresentasse um resultado falso negativo. Assim, o  $EQM$  seria modificado e esta modificação se encontra na equação 4.1.

$$EQM_{modificado}(y) = \frac{1}{2} \sum_{\forall i} [(M - 1)y_i + M + 1](y_i - \hat{y}_i)^2 \quad (4.1)$$

Que pode ser desmembrada em 4.2:

$$EQM_{modificado}(y) = \begin{cases} \mathbf{M} \sum_{i=1}^n (y_i - \hat{y}_i)^2 & , se \quad y = 1 \\ \sum_{i=1}^n (y_i - \hat{y}_i)^2 & , se \quad y = -1 \end{cases} \quad (4.2)$$

Onde  $y_i$  é o  $i$ -ésimo valor da variável status transformada em numérica, podendo ser -1 - bovino não infectado com tuberculose; 1 - bovino infectado com tuberculose;

$M$  é o peso que será imposto para penalizar erros de falso-negativos e, este pode assumir qualquer valor real;

$\hat{y}_i$  representa a  $i$ -ésima saída da rede neural;

$EQM_{modificado}$  é o erro quadrático médio modificado com intuito de penalizar com maior rigor erros falso negativos.

A função indicadora e as modificações na função objetivo do algoritmo *resilient propagation* foram realizadas de forma genérica utilizando o parâmetro  $M$ . Em seguida foram examinados os efeitos dos diversos valores para o parâmetro  $M$  que se encontram na Tabela 4.3.

Tabela 4.3: Valores do Parâmetro M utilizados na ponderação da função indicadora 4.1

Projeto de rede	Valor de $M$
1	1
2	1,5
3	2
4	2,5
5	3
6	3,5
7	4

Assim, foram implementados modelos neurais do tipo MLP para cada um dos 7 valores de  $M$  propostos, a arquitetura destas redes possuíam uma camada escondida com três neurônios utilizando função tangente hiperbólica e uma camada de saída com um neurônio, também, com função tangente hiperbólica. Ainda, foram realizados 100 inicializações com pesos sinápticos diferentes e utilizando o algoritmo *resilient propagation* e, com a função erro quadrado médio modificado ( $EQM_{modificado}$ ) para seu treinamento por batelada e utilizando 2.000 épocas.

Cabe lembrar que para esta rede neural com treinamento utilizando  $EQM_{modificado}$  não se utilizou conjunto teste e nem conjunto validação, pois se bus-

cava uma rede neural que aumentasse a sensibilidade, com a menor perda possível de especificidade. Pode-se observar também que para o valor de  $M=1$  temos o  $EQM_{modificado}$  é igual ao  $EQM$

Assim, a decisão da melhor rede neural, com a aprendizagem modificada foi realizada através de duas etapas: menor  $EQM_{modificado}$  para selecionar dentre as 100 inicializações de pesos sinápticos diferentes, e a segunda etapa foi realizada de forma heurística escolhendo o projeto de rede neural que melhor se adequou ao objetivo imposto ao parâmetro  $M$ . Para essa conclusão, os resultados de sensibilidade e especificidade de cada uma das redes neurais modificada pelo parâmetros  $M$  propostos.

Adicionalmente, neste trabalho, também, foi empreendido após o treinamento dos projetos de rede neural com  $EQM_{modificado}$  pequenas alterações na polarização do neurônio de saída adicionando um valor  $W$  que, por sua vez, teria o efeito de aumentar ou diminuir a entrada da função de ativação, dependendo se ele é positivo ou negativo.

Essas polarizações foram impostas apenas para a rede neural com o valor de  $M=1$ , e a escolha destas alterações foram motivadas também, visando o mesmo propósito dos modelos neurais com  $EQM_{modificado}$ . Deste modo, a Tabela 4.4 mostra os valores de  $W$  utilizados para a construção das redes neurais com a polarização diferenciada. Vale observar que estes valores dependem da estrutura da rede neural e foram estabelecidos experimentalmente.



Tabela 4.4: Tabela de variações  $W$  na polarização da rede neural com parâmetro  $M = 1$

n	variações $W$ nas polarizações
1	0,00
2	0,02
3	0,04
4	0,06
5	0,08
6	0,1
7	0,12
8	0,14

Além disto, foram implementados os procedimentos de validação cruzada do tipo *multifold-cross-validation* que avaliou a boa qualidade de generalização dos modelos discriminantes e das redes neurais. Assim, o conjunto de dados foi dividido em 10 blocos : 9 deles com 12 bovinos e 1 bloco com 8 bovino. Cada um dos 9 blocos de 12 bovinos tinham 6 com status bovino com tuberculose e 6 com status bovino sem tuberculose. Após esta divisão foram selecionados 7 blocos pra treinamento e 3 blocos para teste de 30 formas diferentes. Por fim, foram estimados acurácia, sensibilidade e especificidade com seus respectivos erros amostrais.

Ainda, para a validação cruzada das redes neurais foram utilizados 7 blocos pra treinamento, 2 para teste e 1 para validação.

Finalmente, a escolha do software utilizado foi um passo importante, devendo verificar suas características quanto às necessidades das análises e sua facilidade de uso. Para a construção dos modelos com uso das técnicas de análise discriminante, optou-se por trabalhar com R 3.1.0. e para o treinamento e teste das redes neurais, utilizou-se o MatLab indicado por ser software com pacotes específico para o desenvolvimento desta técnica.

# Capítulo 5

## Resultados

Este capítulo apresenta os resultados obtidos utilizando cada um dos processos de classificação e avaliação do diagnóstico da tuberculose bovina descritas na metodologia, acrescidos de considerações relativas à sua execução; à saber: discriminante linear, discriminante quadrático, rede neural “clássica”, rede neural com ponderação variável e rede neural com polarização variável.

A Figura 5.1 apresenta o plano  $\Delta Bov$  versus  $\Delta Bov - \Delta Av$  onde fica evidenciada graficamente a dificuldade de classificação por qualquer destes dois parâmetros, ou mesmo pelos dois simultaneamente.

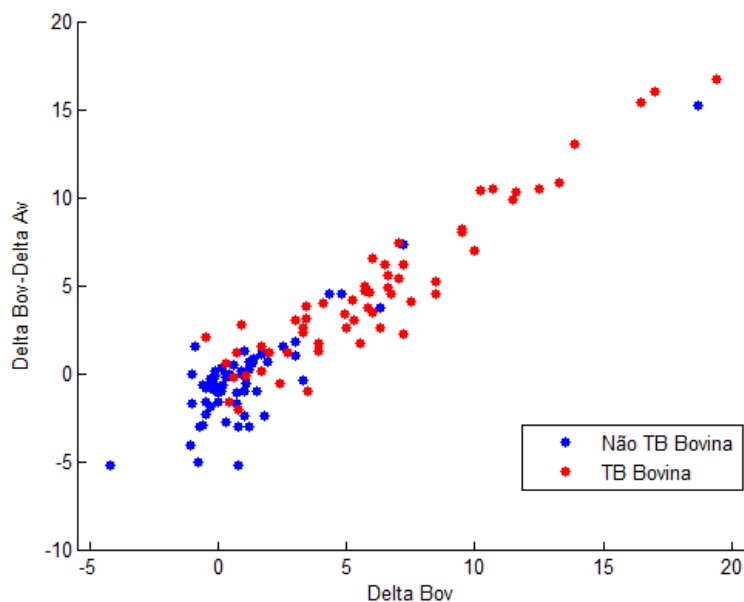


Figura 5.1: Plano  $\Delta Bov$  versus  $\Delta Bov - \Delta Av$

A Tabela 5.1 apresenta as frequências dos resultados para cada um dos testes de diagnóstico quanto à tuberculose: Teste Cervical Simples (TCS) e Teste Cervical Comparativo (TCC). Assim, observa-se que para ambos os testes de diagnóstico havia cerca de 13% de resultados inconclusivos e, o Teste Cervical Simples detectou positivo para tuberculose em bovinos aproximadamente 8% a mais do que o Teste Cervical Comparativo.

Tabela 5.1: Resultados dos Testes de diagnósticos TCS e TCC quanto à Tuberculose Bovina

Teste de Diagnóstico	Positivo		Inconclusivo		Negativo	
	n	%	n	%	n	%
Teste Cervical Simples	42	36,21%	15	12,93%	59	50,86%
Teste Cervical Comparativo	33	28,45%	15	12,93%	68	58,62%

## 5.1 Análise dos Testes Cervicais

A primeira análise realizada neste estudo foi avaliar os índices de desempenho (acurácia, sensibilidade e especificidade) de cada um dos testes de diagnósticos utilizando como “padrão ouro” a variável status.

Para ser possível esta primeira avaliação, os resultados inconclusivos dos testes foram considerados como positivo, ou seja, infectados com tuberculose. Portanto, a partir, deste critério foram realizadas as avaliações cruzadas dos resultados dos testes de diagnósticos com a variável status obtendo os índices de desempenho destes testes de diagnóstico.

Na Tabela 5.2 encontram-se esses resultados e pode-se dizer que para este cenário a acurácia, taxa de acerto global, é praticamente similar para ambos os testes.

O teste cervical simples, por sua vez, mostrou-se com maior sensibilidade, cerca de 8% melhor, em relação ao teste cervical comparativo. Este resultado nos indica que o teste Cervical Simples detectou melhor os bovinos infectados. Em compensação o teste Cervical comparativo possui uma especificidade superior a 90%, ou seja, este teste detectou melhor os bovinos sadios.

Tabela 5.2: Índices de desempenho dos Testes de diagnósticos considerando os resultados de bovinos inconclusivos como infectado por Tuberculose

Teste de Diagnóstico	Acurácia	Sensibilidade	Especificidade
Teste Cervical Simples	83,62%	82,76%	84,48%
Teste Cervical Comparativo	82,76%	74,14%	91,38%

Antes de se construir os modelos discriminantes e os projetos de redes neurais, foram construídos as curvas ROC, que é a principal maneira de expressar a relação entre sensibilidade e especificidade de cada um dos testes de diagnóstico, e portanto, auxiliando na avaliação do poder de discriminação de ambos os testes.

A Figura 5.2 mostra as curvas ROC dos resultados da variável  $\Delta Bov$  e da variável  $\Delta Bov - \Delta Av$ . Neste gráfico, pode-se observar que os resultados destas variáveis alcançaram com similar intensidade a extremidade superior esquerda do gráfico. A variável  $\Delta Bov$  obteve uma área sob a curva de 0,889 enquanto a variável  $\Delta Bov - \Delta Av$ , área 0,879, resultados que podem ser avaliados como bom desempenho discriminatório para ambos os testes de diagnósticos.

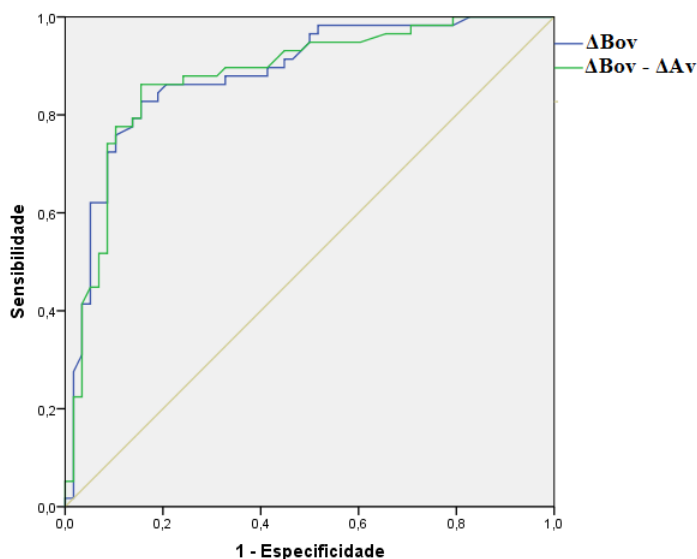


Figura 5.2: Curva ROC dos testes de diagnósticos  $\Delta Bov$  e  $\Delta Bov - \Delta Av$

Para completar a análise do potencial das variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  foram avaliadas sensibilidade, falso alarme (1-especificidade) e índice SP dos pontos de corte obtidos para a construção das curvas ROC.

Cabe lembrar que cada um desses pontos de corte, na verdade, são novas regras de decisão para classificar o bovino em positivo ou negativo para tuberculose bovina.

Sendo assim, a Tabela 5.3 mostra alguns pontos de corte obtidos através da alteração da sensibilidade e do falso alarme dos resultados da variável  $\Delta Bov$ .

Assim, por exemplo, se fosse escolhido o ponto de corte 0,3 haveria  $\Delta Bov > 0,3$  como resultado positivo para tuberculose bovina no teste cervical simples e se  $\Delta Bov \leq 0,3$  então o resultado deste, seria negativo para tuberculose bovina.

Vale destacar na Tabela 5.3 que o ponto de corte de 0,3 no  $\Delta Bov$  resulta uma sensibilidade do diagnóstico da TB bovina no teste cervical simples, que seria cerca de 98%, porém o falso alarme seria elevado também em aproximadamente 52%.

A vista disso, um ponto de corte utilizando um  $\Delta Bov$  inferior a 0,3 seria tão próximo de zero, estabeleceria uma sensibilidade de 100%, porém o índice de falso alarme, também seria de 100% ou muito próximo.

Ainda, pode-se utilizar como ponto de corte o valor que maximizaria o índice SP. Neste caso o ponto de corte seria 1,95 e teria sensibilidade e especificidade próxima a 83% e 85% respectivamente. O interessante deste resultado é que para  $\Delta Bov \geq 2$  o resultado do TCS é inconclusivo ou positivo.

Tabela 5.3: Pontos de corte para a variável  $\Delta Bov$  e seus respectivos índices de sensibilidade, falso alarme e SP

Ponto de Corte	Sensibilidade	Falso Alarme	SP
0,30	98,28%	51,72%	71,04%
0,40	96,55%	50,00%	71,35%
0,50	94,83%	50,00%	70,61%
0,65	93,10%	48,28%	70,89%
0,75	91,38%	44,83%	72,13%
0,85	89,66%	41,38%	73,31%
1,05	87,93%	32,76%	77,24%
1,60	86,21%	20,69%	82,72%
1,70	84,48%	18,97%	82,75%
1,95	82,76%	15,52%	83,62%
2,20	81,03%	15,52%	82,75%
2,60	79,31%	13,79%	82,72%
2,85	77,59%	13,79%	81,84%
3,15	75,86%	10,34%	82,61%
3,30	74,14%	10,34%	81,71%
3,35	72,41%	8,62%	81,62%
3,45	68,97%	8,62%	79,78%
3,70	67,24%	8,62%	78,85%
3,90	65,52%	8,62%	77,91%
4,00	63,79%	8,62%	76,97%

De forma análoga, a Tabela 5.4 descreve novos pontos de cortes criados a partir da variável  $\Delta Bov - \Delta Av$ . Desta maneira se fosse utilizado o ponto de corte -0,60 haveria uma sensibilidade próxima a 95%, porém o falso alarme seria de 50%. Ainda, caso se buscasse um ponto de corte para o TCC que obtivesse uma sensibilidade de 100%, em consequência, teria-se um falso alarme muito próximo de 100%

Se fossem utilizados como ponto de corte da variável  $\Delta Bov - \Delta Av$  que conduz ao máximo SP, este ponto de corte seria 1,15 e teria para o Teste Cervical Comparativo (TCC) a sensibilidade de aproximadamente 86% e especificidade de 84%, ao passo que, se fosse utilizado o mesmo valor de ponto de corte do  $\Delta Bov$ , 1,95 a sensibilidade do teste cervical comparativo seria reduzida para menos de 75% enquanto a especificidade aumentaria para próximo de 92%. Estes valores corroboram o fato do teste comparativo ser mais específico que o o teste simples.

Ainda, caso se utilizássemos um ponto de corte de 1,150, ou seja uma nova regra de decisão para o TCC, haveria rendimentos dos índices de acurácia majorados em 3% se comparados aos resultados do TCC considerando os inconclusivos como positivos.

Tabela 5.4: Pontos de corte para para a variável  $\Delta Bov - \Delta Av$  e seus respectivos índices de sensibilidade, falso alarme e SP

Ponto de Corte	Sensibilidade	Falso Alarme	SP
-0,60	94,83%	50,00%	70,61%
-0,25	93,10%	44,83%	72,89%
-0,15	91,38%	43,10%	73,11%
0,10	89,66%	32,76%	78,04%
0,60	87,93%	24,14%	81,79%
1,15	86,21%	15,52%	85,34%
1,20	84,48%	15,52%	84,48%
1,25	81,03%	15,52%	82,75%
1,40	79,31%	13,79%	82,72%
1,60	77,59%	10,34%	83,51%
1,70	75,86%	10,34%	82,61%
1,95	74,14%	8,62%	82,53%
2,15	72,41%	8,62%	81,62%
2,45	68,97%	8,62%	79,78%
2,70	63,79%	8,62%	76,97%
2,90	62,07%	8,62%	76,01%
3,75	51,72%	6,90%	70,89%
3,90	50,00%	6,90%	69,87%

Após o estudo realizado com as variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  e com os resultados dos teste cervical simples e teste cervical comparativo foram implementados os modelos estatísticos e as redes neurais a partir da combinação das variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  cujo propósito era primeiro aumentar a acurácia do diagnóstico buscando valores próximos a 100%, e caso não fosse obtido sucesso, buscar modelos que potencializassem a sensibilidade, visto que na determinação dos novos pontos de corte para os testes cervicais que aumentassem a sensibilidade, a performance destes, não se mostrou satisfatória, pois o falso alarme também foi elevado.

## 5.2 Modelos Discriminantes

O modelo discriminante foi o primeiro modelo implementado combinando os resultados das variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  para aumentar a acurácia da Tuberculose Bovina. Para esta modelagem, teve-se que verificar os pressupostos de normalidade

das variáveis independentes e a homogeneidade da matriz de variância.

A Tabela 5.5 refere-se aos resultados dos testes de normalidade das variáveis independentes. Nota-se que tanto para a variável  $\Delta Bov$  quanto para a variável  $\Delta Bov - \Delta Av$  o p-valor foi superior a 0,05 e, portanto, não se rejeita  $H_0$ , supondo assim, que os resultados destas variáveis possuam aderência à curva normal. Conclui-se então que não houve violação do pressuposto de normalidade das variáveis.

Tabela 5.5: Resultados dos teste de normalidade Komolgorov-Smirnov para as variáveis independentes

Variáveis independentes	Média	Kolmogorov-Smirnov Z	p-valor
$\Delta Bov$	3,64	1,70	0,19
$\Delta Bov - \Delta Av$	2,24	1,36	0,31

Após avaliar o pressuposto da normalidade das variáveis independentes, falta ainda, investigar a homogeneidade das matrizes de covariância para identificar qual função discriminante apresentará resultados estatisticamente mais consistentes. Desta forma, são apresentados os resultados do teste Box's M na Tabela 5.6 e cujo p-valor foi superior a 0,05, logo não se rejeita a hipótese que as matrizes de covariâncias são homogêneas.

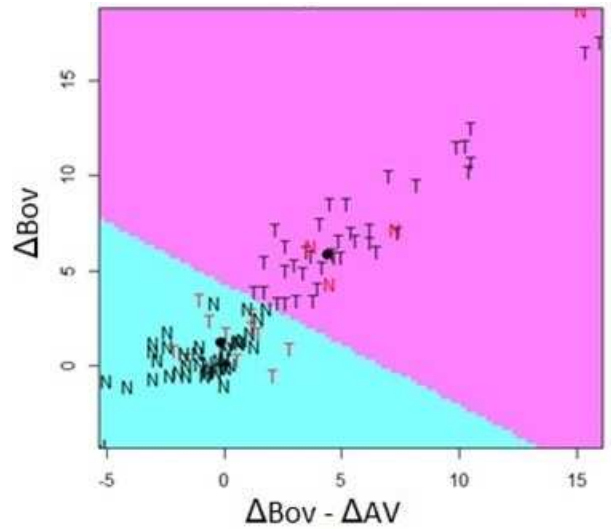
Tabela 5.6: Resultados do teste Box's M de homogeneidade das matrizes de covariâncias das variáveis independentes

Estatística Box's M	p-valor
8,82	0,34

Baseado na análise dos resultados dos testes de normalidade e homogeneidade das variáveis independentes é notório que o modelo discriminante linear é válido, assim foram construídos os modelos linear e quadrático.

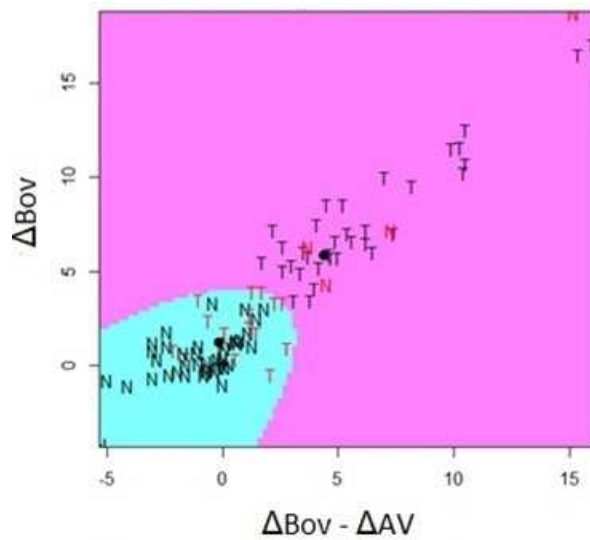


A Figura 5.3 mostra as regiões dos separadores linear e quadrático. Nela pode-se observar que a região quadrática ao classificar bovinos não infectados é bem inferior e que portanto, acaba tendo um número superior de falsos positivos reduzindo a especificidade. Por outro lado a região de infectados e a região de não infectados do modelo linear parece estar em melhor conformidade que as regiões do modelo quadrático.



N - verdadeiro Negativo    T - Verdadeiro Positivo  
 N - Falso Negativo        T - Falso Positivo

(a)



N - verdadeiro Negativo    T - Verdadeiro Positivo  
 N - Falso Negativo        T - Falso Positivo

(b)

Figura 5.3: Separador Discriminante Linear (a) e Discriminante Quadrático (b)

A Tabela 5.7 expressa as medidas de acurácia, sensibilidade e especificidade dos

modelos discriminante linear e quadrático tanto para o conjunto treino quanto para o conjunto teste. Em relação ao conjunto treino, nota-se que o discriminante linear foi mais específico (81,4%) que o modelo quadrático (73,7%), porém em relação a sensibilidade os 2 modelos apresentaram o mesmo índice (94,6%), indicando que não houve diferença significativa na detecção de bovinos doentes. O modelo discriminante linear por ser mais específico, ou seja, detectar bovinos sadios, obteve uma acurácia superior ao modelo quadrático em cerca de 5%.

O conjunto de teste, ou seja, bovinos que não entraram no processo de estimação dos modelos linear e quadrático, mostrou um poder preditivo inferior ao conjunto treino sendo a acurácia do conjunto teste do modelo linear cerca de 81% e do modelo quadrático em torno de 77,3%. A sensibilidade e a especificidade também caíram, o modelo linear apresentou 85% de sensibilidade no conjunto teste e 75,7% de especificidade, enquanto o modelo discriminante quadrático apresentou 80% e 71,3% de sensibilidade e especificidade respectivamente

Estes resultados corroboram o que foi apresentado na Figura 5.3 e nos teste de homogeneidade da matriz de variância, que não rejeitou a homogeneidade das matrizes de variância, porém deve-se analisar, ainda, os resultados da validação cruzada para se ter clareza quanto a boa qualidade de generalização do modelo discriminante.

Tabela 5.7: Índices de desempenho na classificação de Tuberculose Bovina pelos Modelos Discriminante Linear e Quadrático para os conjuntos treino e teste

Modelo Discriminante	Modelo			Teste		
	Acurácia	Sensibilidade	Especificidade	Acurácia	Sensibilidade	Especificidade
Discriminante Linear	86,52%	94,63%	81,47%	81,47%	85,00%	75,72%
Discriminante Quadrático	81,92%	94,66%	73,79%	77,35%	80%	71,31%

Na tabela 5.8 são apresentadas as taxas médias de acurácia, sensibilidade e especificidade, além, dos respectivos erros amostrais das funções discriminantes linear e quadrática, calculadas utilizando validação cruzada do tipo *multifold-cross-validation* com 30 subconjuntos diferentes para treino e teste. Pode-se observar que de uma forma geral, as estimativas médias de acurácia, sensibilidade e especificidade do modelo linear são superiores aos do modelo quadrático, mesmo se considerar estas estimativas dentro do intervalo do erro amostral.

Pode-se dizer que os dois modelos discriminantes possuem uma boa capacidade de generalização, visto que seus erros amostrais são relativamente baixos e variam de 1,16% a 2,98%.

Após a validação cruzada realizada para os modelos linear e quadrático, nota-se que o modelo linear construído com as variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  mostrou

melhor desempenho ao discriminar bovinos quanto à Tuberculose Bovina.

Ao analisar o conjunto teste, o modelo linear possui uma acurácia estimada entre 79,28% a 82,58%; sensibilidade entre 80% a 84,5% e especificidade entre 72% a 76,5%. Estes índices são próximos aos encontrados pelos testes de diagnóstico (TCS e TCC).

Adicionalmente, pode-se dizer que os modelos discriminantes, após os resultados da validação cruzada, apresentaram acurácia próxima aos dos testes de diagnósticos (TCS e TCC), indicando que estes modelos não foram capazes de aumentar a acurácia a valores próximos do 100%.

Tabela 5.8: Validação Cruzada dos modelos discriminante Linear e Quadrático

Modelos Discriminante	Medidas	Treino		Teste	
		Média	± Erro amostral	Média	± Erro amostral
Linear	Acurácia	84,90%	1,92%	80,93%	1,65%
	Sensibilidade	91,90%	1,22%	82,22%	2,24%
	Especificidade	84,52%	2,28%	74,81%	1,86%
Quadrático	Acurácia	81,15%	1,84%	74,81%	2,05%
	Sensibilidade	91,43%	1,28%	80,93%	2,98%
	Especificidade	74,76%	1,71%	71,85%	1,16%

### 5.3 Resultados da rede neural clássica

Após analisar os modelos discriminantes, foram construídos modelos neurais, e a Tabela 5.9 apresenta os resultados obtidos para cada um dos subconjuntos (treino, teste e validação) destes modelos neurais utilizando as variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  e as topologias diferentes que foram propostas neste trabalho.

Pode-se concluir que o classificador MLP com aprendizagem *Resilient Propagation* e com 3 neurônios na camada intermediária foi o projeto de rede neural que apresentou resultados mais satisfatórios dentre todas as topologias avaliadas.

Essa rede, com 3 neurônios na camada intermediária, apresentou menor EQM dentre todos os projetos. Além disto, obteve 87,5% de acurácia no treino e 83,3% no conjunto de teste, ao passo que, ao inserir mais dois neurônios na camada intermediária não houve ganho observado na acurácia do conjunto teste.

Tabela 5.9: Resultados dos Projetos de Rede neurais utilizando topologias diferentes

Topologia	EQM	Acurácia			Sensibilidade			Especificidade		
		Treino	Teste	Validação	Treino	Teste	Validação	Treino	Teste	Validação
1	0,49	82,53%	75,00%	81,81%	81,25%	66,71%	83,34%	77,58%	83,32%	83,32%
2	0,48	83,84%	75,00%	83,35%	90,00%	66,71%	78,64%	90,00%	83,34%	91,71%
3	0,46	87,50%	83,35%	83,35%	91,72%	83,35%	78,6%	84,12%	83,35%	90,00%
5	0,47	87,50%	83,35%	91,72%	91,72%	83,35%	91,72%	84,12%	83,35%	91,71%

A Figura 5.4 mostra a dispersão das saídas da rede neural com 3 neurônios na camada escondida, observa-se que a massa de dados concentrou-se nos extremos, ou seja, esta rede neural possui um bom rendimento, visto que se os dados se concentrasse no limiar da regra de decisão, a rede teria um potencial menor de classificar os bovinos quanto ao diagnóstico de tuberculose.

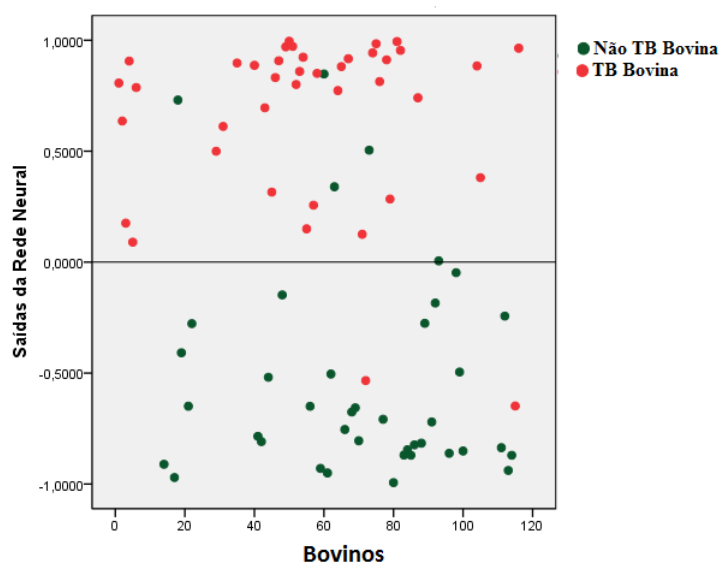


Figura 5.4: Saídas da Rede neural com 3 neurônios

Com a intenção de avaliar a boa generalização do projeto de rede neural que utilizou 3 neurônios na camada intermediária foi implementada a validação cruzada com 30 subconjuntos de treino, teste e validação, e os resultados consolidados encontram-se na Tabela 5.10.

Pode-se notar que o conjunto de teste apresentou estimativas de acurácia que variam entre 79,5% a 84,5%, enquanto a sensibilidade é da magnitude de 79% a 85% e a especificidade entre 79% e 84%.

Ainda, cabe ressaltar que os erros amostrais dos conjuntos treino e teste estão entre 1,5% - 3,0% que podem ser considerados baixos para as estimativas de acurácia, sensibilidade e especificidade, porém estes erros se elevam para 4% quando se estima

sensibilidade e especificidade no conjunto validação, resultado esperado já que este conjunto possui menos dados.

Assim, por possuir erros amostrais relativamente baixos, estes resultados indicam que a rede neural com 3 neurônios na camada intermediária possui boa capacidade de generalização.

Os resultados das estimativas oriundas da validação cruzada realizada no conjunto teste para rede neural com 3 neurônios na camada intermediária apresentam uma performance preditiva similar quando comparados aos dos modelos discriminantes e aos dos testes de diagnósticos.

Tabela 5.10: Validação Cruzada para Rede neural que utiliza 3 neurônios na camada intermediária

Medidas	Treino		Teste		Validação	
	Média	± Erro amostral	Média	± Erro amostral	Média	± Erro amostral
Acurácia	87,26%	1,62%	82,08%	2,51%	80,56%	2,64%
Sensibilidade	90,32%	1,38%	82,22%	2,90%	73,89%	4,05%
Especificidade	85,16%	2,98%	81,39%	2,31%	85,56%	4,06%

Por fim, pode-se dizer que o modelo neural com 3 neurônios, após os resultados da validação cruzada, também apresentou acurácia próxima aos dos testes de diagnósticos (TCS e TCC), indicando que este modelos não fora capaz aumentar a acurácia a valores próximos do 100%.

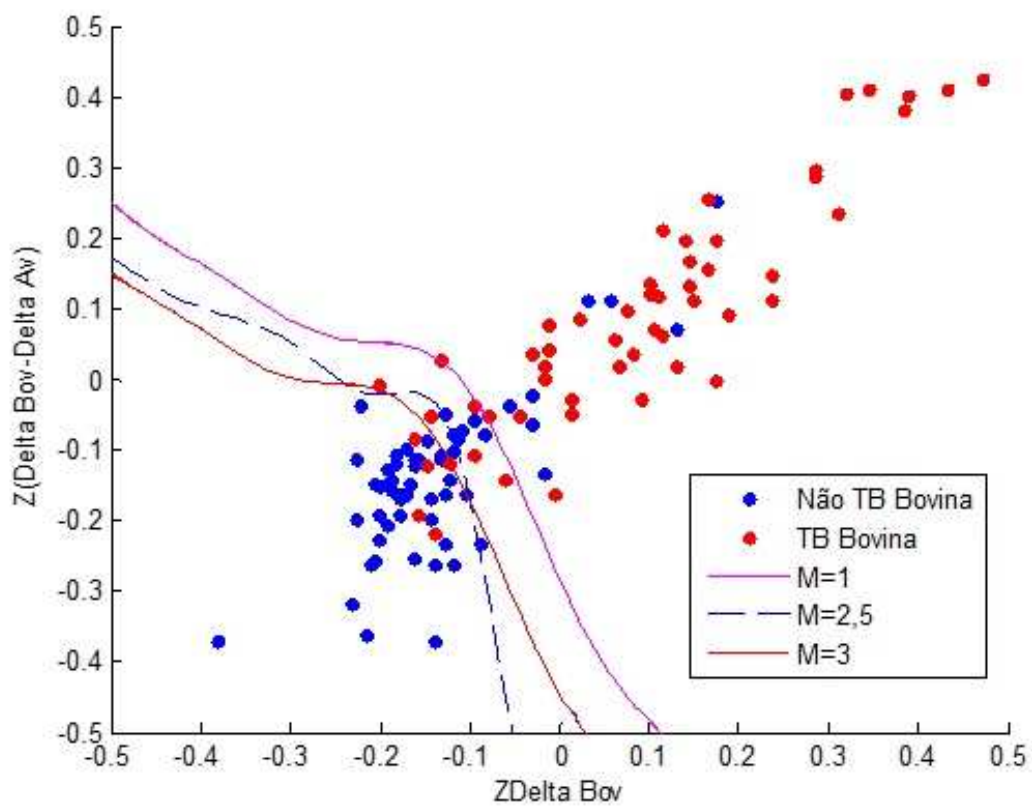
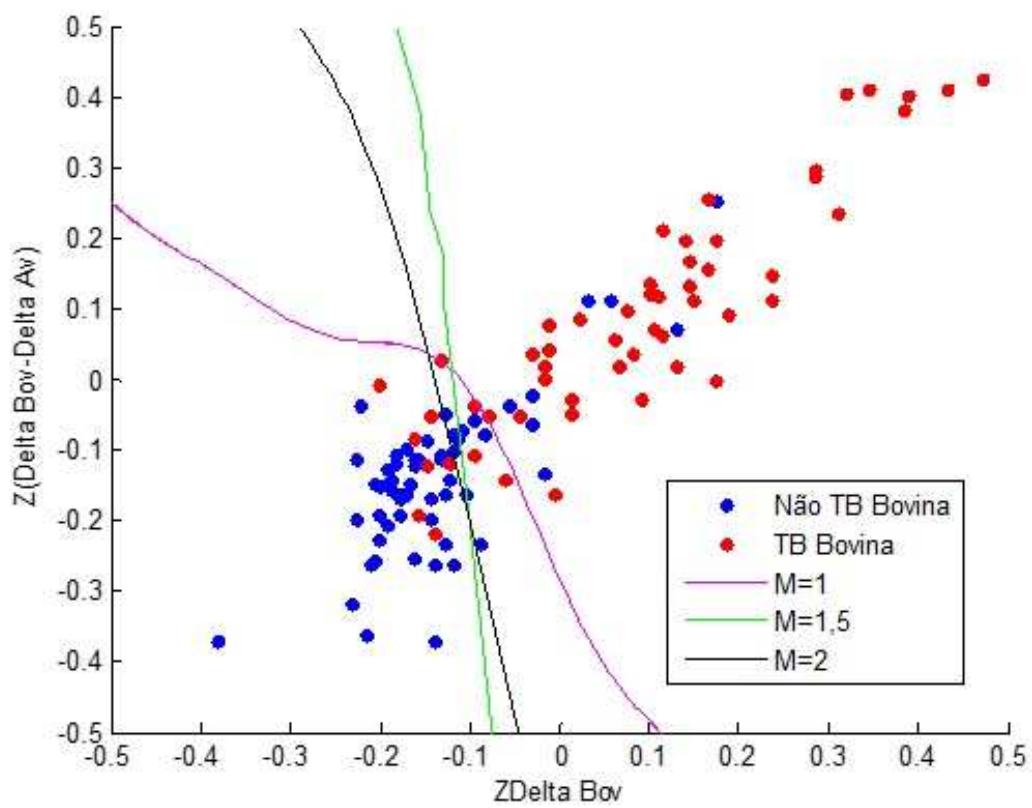
Mediante a estes achados, pode-se dizer que tanto os modelos discriminantes quanto o modelo neural clássico se mostraram ineficientes na busca pelo aumento da acurácia dos diagnósticos da TB bovina.

## 5.4 Resultados da rede neural com Erro Quadrático variável

Á medida que os modelos discriminantes e modelo neural clássico não obtiveram sucesso no aumento da acurácia no diagnóstico da TB bovina, optou-se por buscar modelos que aumentassem a sensibilidade de forma que a queda da especificidade não fosse muito intensa. Já foi observado que buscar novas regras de decisão para os testes cervicais se mostraram ineficazes, pois apesar de obter uma sensibilidade muito elevada este método reduz bastante a especificidade. Assim, decidiu-se por utilizar uma modificação no Erro Quadrático Médio durante o aprendizado das redes MLP com 3 neurônios. Esta modificação propõe uma penalidade,  $M \geq 1$ , que seria imposta quando a rede neural produzisse um erro falso negativo, enquanto um falso

positivo tem penalidade  $M=1$ . Esta modificação já foi apresentada no capítulo de metodologia, bem como uma Tabela com 6 valores para o parâmetro  $M$ .

Ainda, vale ressaltar que para construção dos separadores foi utilizado todo o conjunto de dados no treinamento de redes neurais. Assim, a Figura 5.5 apresenta os separadores que foram produzidas pelo treinamento destas redes neurais para cada valor do parâmetro  $M$ . É notório que ao passo que o valor  $M$  aumenta os separadores tendem a reduzir o número de falsos negativos e em consequência, o número de falsos positivos aumenta, porém de forma menos abrupta que nos pontos de corte dos testes cervicais.



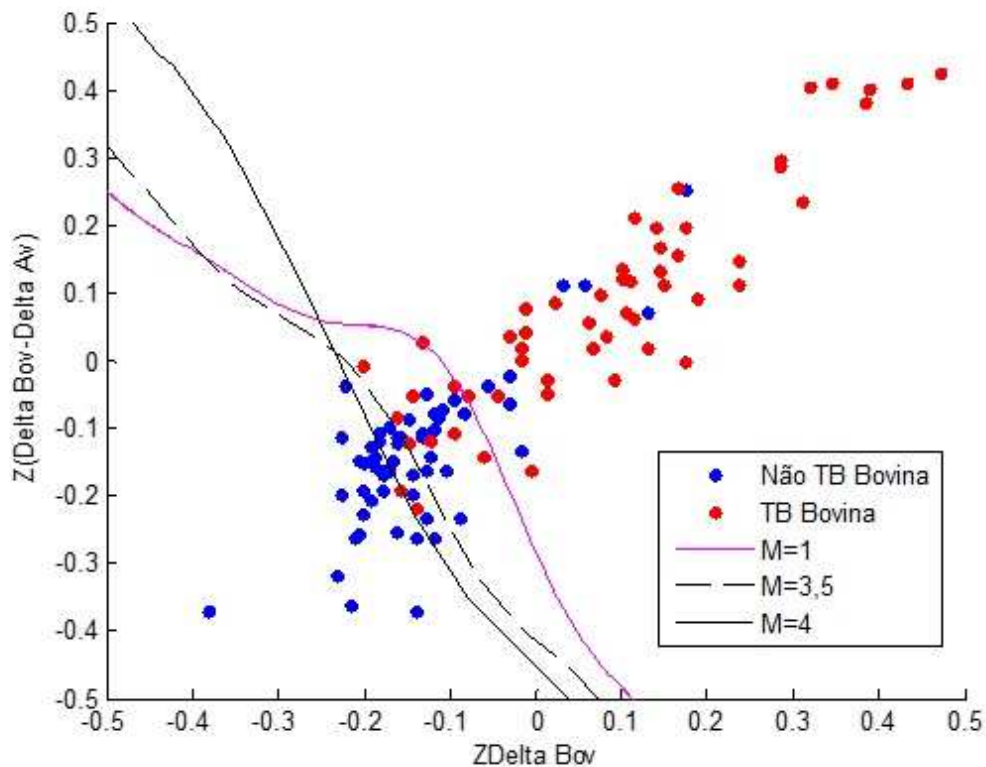


Figura 5.5: Representação dos separadores das redes neurais projetadas para cada valor do parâmetro M

A Tabela 5.11 apresenta o desempenho das redes neurais com o  $EQM_{\text{modificado}}$  por cada valor de M. Nota-se que para o valor de M=4, a rede neural obteve 100% de sensibilidade, porém sua especificidade é muito baixa, cerca de 52%, ocasionando uma acurácia de cerca de 76%.

Apesar de não possuir sensibilidade de 100%, a rede treinada com valor de M=3,5 apresenta uma sensibilidade inferior a 100%, aproximadamente 96%, porém a especificidade se eleva para 62,07%. Este resultado nos indica que o modelo neural com  $EQM_{\text{modificado}}$  é mais adaptável ao problema, visto que a troca de sensibilidade por especificidade é menos abrupta que os pontos de corte dos testes cervicais.



Tabela 5.11: Resultados dos modelos neurais utilizando 3 neurônios na camada intermediária e com as ponderações pelos valores de M na função objetivo

M	VN	FN	VP	FP	Acurácia	Sensibilidade	Especificidade
1,00	50	10	48	8	84,48%	82,76%	86,21%
1,50	46	8	50	12	82,76%	86,21%	79,31%
2,00	43	7	51	15	81,03%	87,93%	74,14%
2,50	44	6	52	14	82,76%	89,66%	75,86%
3,00	42	5	53	16	81,90%	91,38%	72,41%
3,50	36	2	56	22	79,31%	96,55%	62,07%
4,00	30	0	58	28	75,86%	100,00%	51,72%

Mediante a este resultado foi implementado, também, uma validação cruzada para a rede neural que utiliza para aprendizado o  $EQM_{modificado}$  pelo parâmetro  $M=3,5$ . Assim, a Tabela 5.12 mostra as estimativas de média e os erros amostrais para a acurácia, sensibilidade e especificidade. Pode-se dizer que esta rede possui uma boa generalização pois os erros amostrais são baixos, principalmente no conjunto teste, cujos erros amostrais variaram entre 0,75% e 1,5%.

Em relação, aos resultados do conjunto teste a acurácia estimada está entre 77,8% a 80,8% , porém a sensibilidade fica entre 95,9% a 96,8% e especificidade entre 59,8% a 62,8%.

Ainda, a especificidade destes modelo, com  $M=3,5$ , se comparado com as propostas de novas regras de decisão para os testes de diagnósticos que maximizavam a sensibilidade foi cerca de 10% superior. Este resultado nos indica que o modelo neural com  $EQM_{modificado}$  é mais adaptável a busca de maximizar a sensibilidade, visto que a troca de sensibilidade por especificidade é menos abrupta que os pontos de corte dos testes cervicais.

Tabela 5.12: Validação Cruzada para rede neural que utiliza 3 neurônios na camada intermediária e  $EQM_{modificado}$  com parâmetro  $M=3,5$

Medidas	Treino		Teste		Validação	
	Média	± Erro amostral	Média	± Erro amostral	Média	± Erro amostral
Acurácia	79,15%	1,37%	78,97%	0,75%	79,72%	1,50%
Sensibilidade	96,56%	0,62%	95,89%	0,76%	95,94%	2,28%
Especificidade	63,83%	1,14%	61,28%	1,50%	63,89%	3,01%

## 5.5 Resultados da rede neural com polarização diferenciada

Como uma última metodologia trabalhou-se com a polarização diferenciada cujos resultados se encontram na Tabela 5.13 que mostram os efeitos no desempenho da

Rede Neural MLP com 3 neurônios na camada intermediária e com  $M=1$  ao sofrer as variações na polarização do neurônio de saída como proposto na metodologia.

Em comparação aos projetos de redes neurais com  $EQM_{modificado}$  pode-se comparar o desempenho da rede neural modificada com parâmetro  $M=4$  com a rede clássica e com polarização *bias* de  $W = 0,14$ , pois ambas têm 100% de sensibilidade porém a especificidade do modelo neural com polarização positiva é inferior a 1,5%.

Ao confrontar o modelo neural modificado com  $M = 3,5$  com a rede neural clássica e com polarização *bias* de  $W = 0,12$  estas duas redes teriam a mesma sensibilidade, porém a especificidade da rede com  $M = 3,5$  seria superior em torno de 5% em relação a outra rede. tais resultados nos indica que o é menos relevante que modelo neural com erro modificado na busca por aumentar a sensibilidade, pois também como nos pontos de corte a especificidade se reduz de forma mais severa.

Tabela 5.13: Resultados das polarizações *bias* nos modelos neurais utilizando 3 neurônios na camada intermediária e sem ponderação na função objetivo

$W$	VN	FN	VP	FP	Acurácia	Sensibilidade	Especificidade
0,00	50	10	48	8	84,48%	82,76%	86,21%
0,02	48	7	51	10	85,34%	87,93%	82,76%
0,04	45	6	52	13	83,62%	89,66%	77,59%
0,06	43	4	54	15	83,62%	93,10%	74,14%
0,08	37	3	55	21	79,31%	94,83%	63,79%
0,10	35	2	56	23	78,45%	96,55%	60,34%
0,12	33	2	56	25	76,72%	96,55%	56,90%
0,14	29	0	58	29	75,00%	100,00%	50,00%

## 5.6 Comparação entre os Testes cervicais e os principais modelos

Foram construídas curvas ROC para comparar o poder discriminatório das variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  com os modelos que se mostraram mais adequados neste trabalho. Assim, pela Figura 5.6, pôde-se avaliar que as variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  separadamente tem um poder de classificação quanto à tuberculose bovina parecidos com o modelo discriminante linear, porém em relação à rede neural clássica com 3 neurônios na camada intermediária nota-se uma pequena vantagem no poder classificatório deste último modelo por possuir a curva mais superior a esquerda. Ainda, a área sob a curva do modelo neural é a única que é superior a 0,9 e pode ser classificada com excelente desempenho.

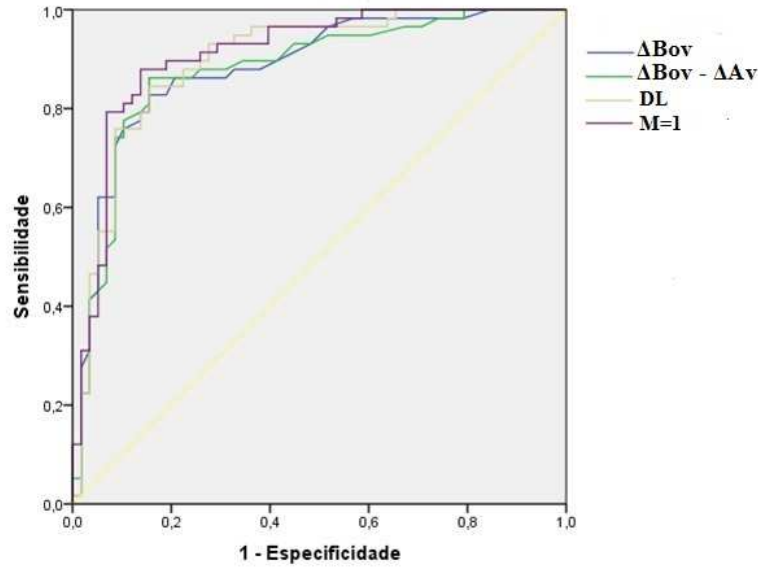


Figura 5.6: Curvas ROC das variáveis  $\Delta B_{ov}$  e  $\Delta B_{ov} - \Delta A_v$  e dos modelo discriminante linear e modelo neural clássico com 3 neurônios na camada intermediária

A Figura 5.7 mostra que as curvas ROC do modelo neural modificado por  $M = 4$  e da rede neural com polarização diferenciada  $W = 0,14$  atingem praticamente a sensibilidade máxima com o mesmo índice de falso alarme (1-especificidade), ou seja, as duas curvas ao atingir ordenada 1 (sensibilidade 100%) estão localizadas na abscissa 0,5 (Falso alarme 50%). Ainda, nota-se que a rede neural com  $M = 4$  possui maior área sob a curva indicando ser melhor modelo classificatório quanto ao diagnóstico de Tuberculose e, inclusive, em relação as variáveis  $\Delta B_{ov}$  e  $\Delta B_{ov} - \Delta A_v$ .

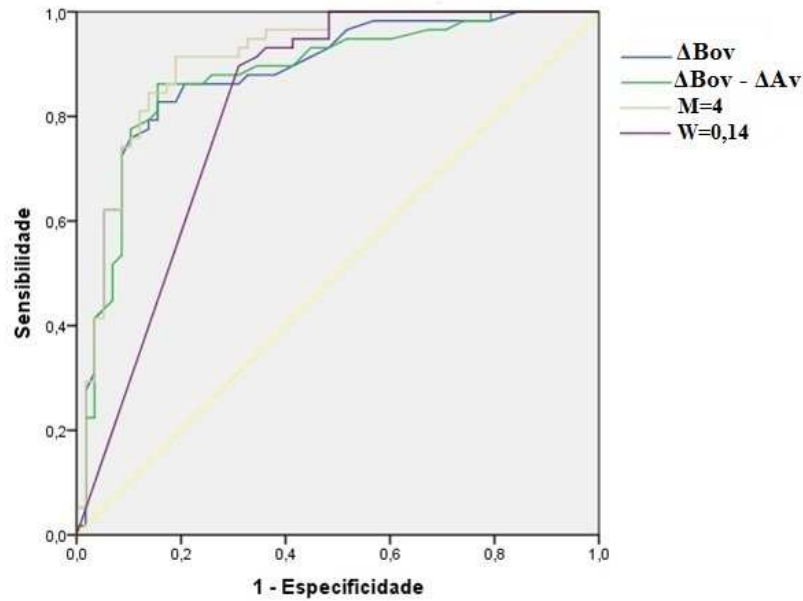


Figura 5.7: Curvas ROC das variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$ , modelo neural modificado por  $M = 4$  e modelo neural clássico com polarização positiva de  $W = 0,14$

A Figura 5.8 mostra que as curva ROC da rede neural com  $M = 3,5$  e da rede neural com  $M = 1$  e  $W = 0,12$  possuem maior poder classificatório em relação as variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$ . Ainda, se comparados os dois modelos de rede neural não se pode afirmar que algum destes leva vantagem no potencial de diagnosticar bovinos com tuberculose.

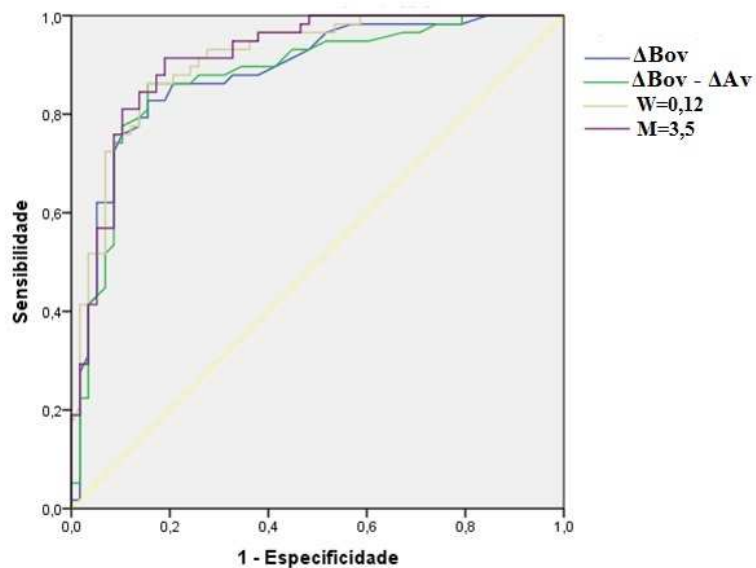


Figura 5.8: Curvas ROC das variáveis  $\Delta Bov$  e  $\Delta Bov - \Delta Av$ , modelo neural modificado por  $M = 3,5$  e modelo neural clássico com polarização positiva de  $W = 0,12$

# Capítulo 6

## Conclusão

### 6.1 Discussão

O desenvolvimento da biologia molecular e a melhora na compreensão do conceito de imunidade acarretaram no surgimento de diversos estudos que, ou desenvolveram novos métodos para o diagnóstico da tuberculose bovina, ou, simplesmente, realizaram estudos de sua eficiência e praticidade ao serem aplicados [9].

Segundo o MAPA, sobre os métodos de diagnósticos pode-se classificar em métodos indiretos que são os testes sorológicos detectam alterações dos componentes séricos - e métodos de avaliação da resposta imunológica celular [9].

Já os métodos diretos são caracterizados pelos métodos moleculares denominados reação da polimerase em cadeia (PCR), que tem como princípio básico a detecção de um fragmento de DNA específico do gênero [9].

Em 2007, Ruggiero et al. fizeram uma revisão dos principais métodos de diagnósticos conhecidos àquela época: teste clínico, testes tuberculínicos, cultura, exame *post-mortem*, ELISA,  $\gamma$ -interferon e ensaios moleculares como RFLP e em especial PCR. Eles concluíram que apesar dessa numerosa quantidade de métodos de diagnóstico para a tuberculose bovina; ainda não se tinha encontrado um método que fosse ao mesmo tempo rápido, eficiente e com baixa complexidade para substituir os testes tuberculínicos que são os preconizados pelo PNCEBT.

Os métodos de diagnósticos da TB bovina podem ser separados em *ante mortem* que são realizados com o animal vivo; e os métodos de inspeção *post mortem* que tem por objetivo identificar o agente através de análise de lesões suspeitas. Portanto, os exames *ante mortem* são métodos de diagnósticos de investigação, enquanto os métodos de inspeção *post mortem* possuem um caráter confirmatório [9].

Em relação aos métodos de diagnósticos *ante mortem* preconizados para bovinos, Ruggiero et al [58] concluíram que o exame clínico apresentava complexidade na sua utilização devido aos sintomas ocorrerem apenas em estágios avançados da doença.

Outro método preconizado pelo PNCEBT avaliados por eles, foram os testes tuberculínicos, que apesar de possuir restrições quanto a sensibilidade e especificidade, ainda eram os eleitos devido a sua simplicidade.

O MAPA [9], em 2006, considerava  $\gamma$ -interferon útil para detectar animais tuberculosos, e que estes ensaios já vinha sendo empregado como teste complementar da tuberculinização em alguns países, como Austrália e Nova Zelândia.

Ruggiero et al. [58] concordavam com a análise do MAPA em relação ao ensaio  $\gamma$ -interferon e, ainda, apresentaram como vantagem o fato de ser um procedimento não invasivo e que podia ser realizado várias vezes sem intervalo de tempo.

Contudo, Abrahão [1], em 1999, tinha observado que o ensaio de  $\gamma$ -interferon, oferecia algumas restrições, como o elevado custo, a necessidade de um tempo para o processamento das amostras de sangue (até 8h após a coleta), além da possibilidade da ocorrência de resultados falso positivos, em razão das reações cruzadas com micobactérias inespecíficas.

Marassi et al. em 2013 [39], realizaram um estudo com diversas técnicas *ante mortem* e obtiveram como resultados: sensibilidade superior a 91% e especificidade em torno de 87% ao utilizarem o ensaio de  $\gamma$ -interferon com amostras coletadas no mesmo dia da execução do TCC; aproximadamente 74% de sensibilidade e 87% de especificidade após 7 dias da execução do TCC e aproximadamente 71% de sensibilidade e 80% de especificidade para as amostras coletadas no dia 21 da realização do TCC.

Marassi et al.[39] também obtiveram sensibilidade e especificidade dos ensaio do ELISA, outra metodologia para diagnóstico da tuberculose *ante mortem*, utilizando o mesmo período da coleta do ensaio  $\gamma$ -interferon. No entanto, os índices de sensibilidade e especificidade foram pífios, cujo maior valor de sensibilidade foi encontrado no ensaio realizado após 7 dias da execução do TCC, 53,8%. Enquanto o maior valor de especificidade foi encontrado com amostras coletadas após 21 dias da execução com 77,7%.

Neste estudo os autores acreditavam que se fosse adotada uma estratégia mais abrangente, ou seja, com utilização de mais teste de diagnósticos além dos Preconizados pelo PNCEBT, esta deveria incluir métodos de diagnóstico que pudessem identificar a resposta imune celular e humoral, uma vez que animais de um mesmo rebanho poderiam se encontrar em diferentes estágios da infecção [39].

Em outro estudo, Marassi et al.[40] avaliaram a detecção de anticorpos específicos em vacas após a injeção de PPD. Sendo o principal objetivo analisar se os testes intradérmicos influenciariam os resultados dos ELISAs que utilizaram as proteínas MPB70 e MPB83 como antígenos de captura em um rebanho naturalmente infectado. Eles concluíram, através de testes estatísticos, que o desempenho de ambos os ELISAs foi comparável. No entanto, os resultados demonstraram que entre os

animais positivos, houve um aumento de reatividade ao MPB70 em comparação ao MPB83.

Numa opinião diferente de Marassi et al. [40] possuíam, Fráguas et al. [25] ao avaliarem o ELISA, julgaram que este teste não deveria ser adotado como complementar, pois sua sensibilidade cai consideravelmente em animais que já passaram pela tuberculinização.

Já Ruggiero et al. [58] complementam que o ELISA poderia ser utilizado como um exame complementar aos ensaios baseados na imunidade celular, pois eram úteis na identificação dos anérgicos.

Num estudo realizado com técnicas moleculares, em 2009, por Figueiredo et al. [23] o m-PCR foi utilizado numa investigação *ante mortem* com o muco nasal dos bovinos. Embora a utilização deste método apresente limitações, já que necessita que o bovino esteja em estágio avançado da infecção para que sejam obtidos o muco nasal, ele se mostrou um método auxiliar eficaz, desde que fosse realizado de forma regular ou, ainda, como um método de confirmação para animais com os resultados dos testes intradérmicos inconclusivos.

Em referência as inspeção *post mortem*, Ruggiero et al. [58] e Pinto et al. [51] ressaltam sua utilidade para a comprovação do diagnóstico e para a avaliação da eficácia das provas de tuberculina.

Cobos-Marín et al. [12] afirmavam, em 2003, que a identificação do agente *M.bovis* em laboratórios, depois do isolamento bacteriano, se baseia em morfologia colonial e testes bioquímicos, contudo esses testes são altamente demorados e requerem laboratórios especializados.

Ruggiero et al. [58] acreditava que as técnicas bacteriológicas existentes em 2007 necessitavam de grandes concentrações de bacilos e que, por isto acabavam consumindo bastante tempo para crescimento em meio de cultura apropriado. Além disto, Zanini et al.[61], em 2001, já alertavam para a complexidade da utilização das técnicas bacteriológicas, pois nem sempre eram atendidas todas as exigências.

Em relação aos testes histopatológicos Ruggiero et al. [58] afirmavam que são testes rápidos e baratos, contudo, também exigiam grandes concentrações de bactérias, além desta técnica poder muitas vezes apresentar falhas na identificação do agente *M.bovis*, pois outros microrganismos podiam produzir lesões semelhantes a da tuberculose bovina.

Ainda sobre exames confirmatórios *post mortem*, Charro e Osório [10] realizaram um estudo anatomopatológico e pesquisaram o bacilos álcool-ácido resistentes (BAAR) em amostras de tecidos provenientes da inspeção de bovinos em matadouros-frigoríficos de Mato Grosso do Sul. Eles observaram que nos cortes histológicos sob coloração de Ziehl-Neelsen-Faraco foram identificados o BAAR resistentes em quatro das seis amostras. Enquanto sob coloração de HE três das seis

amostras apresentaram granuloma sugestivo de tuberculose e três mostraram lesões difusas. O exame histopatológico, sob coloração de HE, é considerado sugestivo de tuberculose, e sob coloração especial de Ziehl-Neelsen-Faraco foi que se puderam evidenciar a presença de BAAR, portanto, eles concluíram que os dois métodos de diagnósticos se complementavam.

Para Ruggiero et al. [58], os ensaios moleculares em especial o PCR nas inspeções *post mortem* utilizando como amostra as lesões suspeitas de TB bovina não substituiria os métodos clássicos de isolamento e identificação do agente, mas, somente, auxiliariam na redução do tempo para a confirmação do diagnóstico, pois apresentavam algumas restrições técnicas que impediam a sua aplicação em 2007. Entre elas: a extração de DNA de amostras clínicas, especialmente em razão da pequena quantidade de bacilos presentes em lesões de bovinos tuberculosos; a escolha de *primers* adequados, que assegurem a especificidade da reação; além do alto custo do diagnóstico.

Rodriguez et al. [54] apresentaram uma combinação para o diagnóstico confirmatório da tuberculose bovina utilizando o isolamento através de bacteriologia clássica, seguido da identificação por métodos moleculares. *Ospoligotyping* é um método rápido e de baixo custo para tipificação e foi utilizado com sucesso em alguns estudos com cepas de *M. bovis*.

Parra et al.[50], em 2008, desenvolveram um teste de PCR em tempo real de alta especificidade, que não reagia a outros microorganismos causadores de linfadenite. A sensibilidade deste método não foram tão relevantes, aproximadamente 61%, quando aplicado em carcaças sem lesões visíveis, e 81% em carcaças com lesões crônicas.

Figueiredo et al.[22] realizaram um estudo, em 2008, de investigação *post mortem* em um rebanho em que foram desenvolvidos sistemas PCR para a detecção de espécies pertencentes a *M. tuberculosis* complexo (MTC). Foram utilizados dois conjuntos de *primers*. Assim, foram isoladas as colônias a partir de lesões suspeitas de tuberculose bovina e testadas pela reação múltipla em cadeia da polimerase (m-PCR). Como resultado deste estudo a m-PCR identificou, com sucesso, 88,24% das colônias isoladas como *M. bovis*.

Em 2012, Furlanetto et al.[27] acompanharam o abate e a inspeção de 41.193 bovinos, sadios ao exame *post mortem*, em sete abatedouros no estado de Mato Grosso com o objetivo de avaliar métodos complementares de diagnóstico (histopatológicos, bacteriológicos e moleculares), no julgamento de lesões suspeitas de tuberculose observadas durante a inspeção *post mortem* de rotina em abatedouros. Neste estudo eles concluíram que as lesões pelo exame macroscópico concordaram em 93% com os resultados obtidos por meio da PCR.

Corcoul et al.[19] realizaram um estudo com um rebanho composto por 5211 animais na França cujo objetivo foi estimar e comparar as sensibilidades e espe-



cificidades dos testes bacteriológicos, histopatológicos e PCR na ausência de um “padrão-ouro” e usando análise de classe latente. Neste estudo foram desenvolvidos modelos bayesianos, permitindo a dependência entre exames bacteriológicos e PCR, enquanto assumiu-se independência dos exames histopatológicos. A sensibilidade da PCR foi maior do que os exames bacteriológicos (em média 87,7% [82,5-92,3%] versus 78,1% [72,9-82,8%]) enquanto que a especificidade de ambos os testes foi muito boa (em média 97,0% para a PCR [94,3-99,0%] e 99,1% para bacteriologia [97,1-100,0%]). O exame histopatológico foi pelo menos tão sensível como PCR (em média, 93,6% [89,9-96,9%]), mas menos específico do que os outros dois testes (em média, 83,3% [78,7-87,6%]). Estes resultados sugerem que a PCR tem potencial para substituir exames bacteriológicos na confirmação de TB bovina em amostras de lesões sugestivas apresentadas a partir de bovinos suspeitos.

Em relação a utilização de técnicas computacionais no auxílio do diagnóstico da TB bovina, foi realizado um trabalho por Gracioso et al. [28], em 2013, cujo objetivo era verificar a eficiência de um Sistema Neuro-*Fuzzy* NEFCLASS (Neuro-*Fuzzy* Classification) na classificação de imagens de tuberculose através dos seus atributos de textura extraídos utilizando-se a Transformada *Wavelet*. Para realização dos testes, foram utilizadas imagens cedidas pelo Instituto Biológico de São Paulo. Os testes realizados para classificação das imagens entre patológicas e não patológicas para imagens de exames de tuberculose apresentaram uma precisão para a imagens pertencentes a classe positivo, o de 75%, sendo. Já para as imagens da classe negativo, o índice de acerto foi de 68,42%. Apesar dos resultados obtidos não serem expressivos, para os autores do estudo estes resultados serviriam de base para o desenvolvimento de sistemas de auxílio ao diagnóstico diminuindo o esforço humano e automatizando toda ou parte da classificação das imagens entre patologias.

Vordermeier et al. [59] desenvolveram um teste cutâneo para tuberculose bovina utilizado para a diferenciação de infecção por *M. Bovis* a animais que foram vacinados, em inglês o nome deste teste é DIVA. Segundo os autores, com o progresso no desenvolvimento de uma vacina contra a TB bovina existia uma necessidade de testes de diagnóstico para diferenciar entre animais infectados e vacinados (DIVA). Entre os resultados, destaca-se que um coquetel de *Mycobacterium tuberculosis* com proteína recombinante do complexo antígenos ESAT-6, CFP-10, MPB70, MPB83 provocou reações do tipo tardio de hipersensibilidade nos testes cutâneos em 78% dos bovinos naturalmente infectados. Outro resultado encontrado pelos autores é que foi demonstrado pela primeira vez a utilidade de péptidos sintéticos como antígenos do teste cutâneo, promissores para tuberculose bovina. Além disto, o coquetel de antígenos proposto não induziu respostas de pele em bovinos vacinados com BCG, apesar deles estarem sensibilizados para fortes respostas de tuberculina.

Conlan et al. [13] realizaram um estudo para verificar os benefícios potenciais

da vacinação como forma suplementar de controle da tuberculose bovina em bovinos, visto que não é usado atualmente dentro de qualquer programa de controle internacional, e é ilegal na EU. Assim, novos testes de diagnóstico que diferenciam infectados de animais vacinados (DIVA) oferecem a possibilidade de introduzir a vacina no âmbito de programas de erradicação existentes. Portanto, o objetivo do trabalho de Conlan et al. foi de explorar a viabilidade da vacinação no gado como forma complementar de controle. Neste estudo foi estimado uma especificidade pelo teste DIVA de pelo menos, 99,85% e sensibilidade um pouco superior a 40%. Ainda, dados a partir de animais infectados experimentalmente sugerem que esta especificidade superior a 99,85% poderia ser conseguida em animais vacinados utilizando um coquetel de três antígenos na composição do teste DIVA mantendo ao mesmo tempo uma sensibilidade de 73.3% (95% CI: 61,9, 82,9%) em relação a detecção *post-mortem*.

Portanto, foi nesta realidade que se inseriu o presente estudo, onde diagnósticos *ante mortem* alternativos propostos até o momento, mostraram-se ineficientes como o ELISA ou promissores como as técnicas moleculares e o  $\gamma$ -interferon. No entanto, o  $\gamma$ -interferon possui restrições quanto ao custo elevado, enquanto as técnicas moleculares atualmente, são utilizadas com sucesso apenas como complementares aos exames preconizados ou em inspeções *post mortem*.

Desse modo, foram propostos métodos de classificação supervisionadas que auxiliassem no diagnóstico da TB bovina através da combinação, ou não, dos resultados do  $\Delta_{\text{bov}}$  e do  $\Delta_{\text{bov}} - \Delta_{\text{av}}$  com dois propósitos distintos: o primeiro seria aumentar a acurácia em relação aos testes cervicais, e o segundo, aumentar a detecção de bovinos doentes (sensibilidade) o máximo possível sem que houvesse um relevante aumento dos falsos positivos (especificidade).

Além disto, vale ressaltar que optou-se por trabalhar apenas com os resultados das Inoculações do PPD bovino e PPD aviário devido a sua simplicidade, pelo seu baixo custo e por ser comum na rotina do produtor. Além é claro de ser as provas tuberculínicas, as preconizadas pelo PNCEBT.

O método preconizado pelo PNCEBT deixa como opção ao produtor se ele quer manter os bovinos com resultados positivos e inconclusivos no TCS vivos e separados dos demais por no mínimo 90 dias até que seja realizada o TCC. somente, em caso de positivos no TCC é que os bovinos seriam sacrificados.

Neste estudo, optou-se por avaliar a acurácia dos testes cervicais separadamente e não seguindo o ciclo estabelecido pelo PNCEBT, pois, queria-se determinar o potencial de cada teste cervical. Assim, os inconclusivos em ambos os testes foram considerados positivos. A acurácia foi de 83,62% e 82,76% para o TCS e o TCC respectivamente.

A Análise Discriminante linear e quadrática foram os dois primeiros modelos

propostos neste trabalho, que combinou  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  com a intenção de aumentar a acurácia do diagnóstico da TB bovina. Porém foram obtidos como resultados as estimativas de acurácia em torno de 80,9% (79,3% a 82,6%) no modelo linear e 74,8% (72,8% a 76,9%) no modelo quadrático. Tais estimativas estão longe do alvo que seria 100% de acurácia. Além disto, tanto o modelo discriminante linear quanto o quadrático não conseguiram obter acurácia superior as encontradas nos testes cervicais.

As redes neurais clássicas do tipo MLP com aprendizado utilizando o algoritmo *Resilient propagation*, foram o segundo método delineado para determinar um diagnóstico da TB bovina com intuito de aumentar a acurácia. Assim, foram obtidos os melhores resultados quando se utilizou 3 neurônios na camada intermediária. A estimativa de acurácia ficou em torno de 82,1% 79,6% a 84,6%, ou seja, também distante da meta que seria 100% de acurácia.

Assim, tanto os modelos discriminantes quanto os modelos neurais clássicos se mostraram ineficientes na construção de um método de diagnóstico da TB bovina que alcançassem acurácias próximas de 100%, ou pelo menos, com maior acurácia que os testes cervicais.

Após o fracasso na tentativa de buscar uma metodologia através dos resultados  $\Delta Bov$  e  $\Delta Bov - \Delta Av$  que diagnosticassem a TB bovina com acurácia próxima a 100%, se buscou métodos que maximizassem a detecção de bovinos doentes, porém sempre observando se ocorria uma queda drástica da especificidade.

Vale ressaltar que, se por um lado é de grande importância maximizar a sensibilidade - pois evita que animais doentes permaneçam no rebanho e contaminem outros animais - sendo essencial para o controle da TB bovina; aceitar uma especificidade muito baixa seria uma estratégia com impacto econômico muito negativo também, principalmente em rebanho que possuam micobactérias ambientais ou em rebanhos de gado leiteiro.

Portanto, como primeira estratégia se buscou um ponto de corte com maior sensibilidade para o teste TCS, destacando-se o valor de  $\Delta Bov > 0,3$ . Com esta nova regra de decisão o TCS teria uma sensibilidade próxima de 98,3%, porém a especificidade seria pífia, inferior a 50%. Ao realizar o mesmo processo para o TCC, distingue-se a regra de decisão:  $\Delta Bov - \Delta Av > -0,6$ . Obteve-se uma sensibilidade próxima a 95%, porém, com especificidade (50%).

É claro que seria possível uma nova regra de decisão para o TCS ou TCC cuja a sensibilidade fosse máxima, porém a especificidade seria menor ainda se aproximando de 0%.

Outra metodologia proposta no trabalho com o objetivo de aumentar sensibilidade, foram os modelos neurais com os  $EQM_{modificado}$  pelo parâmetro M. Este modelo tinha por finalidade penalizar um erro falso negativo com maior intensidade

a um erro falso positivo. Os resultados com maiores sensibilidades desta modelagem ocorreram para valores de  $M=3,5$  e  $M=4$ . Ao utilizar o valor de  $M=4$ , a sensibilidade foi de 100% porém como era esperado, a especificidade reduziu consideravelmente, para cerca de 51%.

Apesar disto, esses resultados são melhores que os encontrados quando se propôs novas regras de decisão para o TCS e TCC, pois a sensibilidade deles foram de 98,3% e 94,8% respectivamente, porém suas especificidade se mantiveram em torno de 50%.

A rede neural modificada usando o valor  $M = 3,5$  obteve valores de sensibilidade próximo de 96,6% e especificidade próximos a 62%. Portanto, este modelo apresentou sensibilidade aproximadamente 2% inferior e especificidade 13% superior quando comparada com a sugerida regra de decisão  $\Delta Bov > 0,3$  para o TCS.

Além disso, se a regra de decisão fosse  $\Delta Bov > 0,4$  para o TCS, o modelo neural com  $M = 3,5$  possuiria uma sensibilidade igual a desta regra, porém com uma especificidade 12% maior.

Quando se compara o modelo  $M = 3,5$  com o modelo  $M = 4$ , o ganho foi de cerca de 11% na especificidade e uma perda de 3,5% na sensibilidade.

Desta forma, o modelo neural com *Erro<sub>m</sub> modificado* se mostrou mais habilitado a maximizar a sensibilidade, uma vez que a redução da especificidade ocorreu de maneira menos severa que nos novos pontos de corte para os testes cervicais.

Acredita-se que isto ocorra, pois mesmo penalizando com maior intensidade os falsos negativos, a função erro do modelo neural modificado também penaliza os falsos positivos, evitando que ocorra uma demasiada mal classificação de bovinos sadios como positivos.

O estudo de polarização diferenciada mostrou que apesar da sensibilidade ser a mesma da rede neural modificada por  $M = 4$ . A rede modificada por  $M = 4$  possui especificidade 5% superior que a rede neural clássica com polarização positiva de  $W = 0,14$ .

De forma análoga, se pode dizer a respeito da rede neural modificada por  $M = 3,5$  em comparação a rede neural clássica com polarização positiva de  $W = 0,12$ . Apesar de possuírem sensibilidade similares, a especificidade da primeira é superior.

## 6.2 Considerações Finais

- Em relação aos modelos proposto de análise discriminante e neural clássico, pode-se concluir que eles foram ineficientes no objetivo de aumentar a acurácia dos diagnósticos de tuberculose bovina.
- Em relação aos modelos propostos com a finalidade de aumentar a sensibili-

dade, sem que houvesse uma redução radical da especificidade. Conclui-se que a rede neural com o *Erro<sub>m</sub>odificado* se mostrou um modelo alternativo de diagnóstico promissor no aumento detecção de bovinos infectados, pois, apesar de priorizar a sensibilidade, a especificidade destes modelos se reduziu de uma forma menos drástica que as regras de decisão propostas para o TCS e o TCC, ou mesmo que os modelos neurais com polarização positiva.

- O problema é complexo devido à baixa acurácia dos testes cervicais, e é dificultada pelo número restrito de amostras disponíveis. A utilização de testes altamente mais eficazes e de dados de populações maiores, caso se tornem disponíveis no futuro, são os caminhos a trilhar para a continuação da pesquisa.
- Por fim, mesmo com as restrições acima mencionadas, acredita-se que este trabalho pôde contribuir para a pesquisa e desenvolvimento do controle da Tuberculose bovina.

# Referências Bibliográficas

- [1] ABRAHÃO, R., 1999, “Tuberculose humana causada pelo *Mycobacterium bovis* considerações gerais e a importância dos reservatórios animais.” *Arch Vet Scienc.*, v. 4, pp. 5–15.
- [2] ANDERSON, T., 2003, *Introduction to multivariate statistical analysis.*, v. 3. New York, Wiley.
- [3] BEER, J., 1988, *Doenças Infecciosas em Animais Domésticos*. São Paulo, Roca.
- [4] BIET, F., BOSCHIROLI, M., THOREL, M., et al., 2005, “Zoonotic aspects of *Mycobacterium bovis* and *Mycobacterium avium*–intracellulare complex MAC”, *Vet Res.*, v. 36, pp. 411–436.
- [5] BOX, G., 1949, “A general distribution theory for a class of likelihood criteria.” *Biometrika*, v. 36, pp. 317–346.
- [6] BRAGA, A., CARVALHO, A., LUDEMIR, T., 2000, *Redes Neurais Artificiais, Teoria e Aplicações*. Rio de Janeiro, Ed. LTC.
- [7] BRASIL, 2001. “Instrução Normativa 2/2001. Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose.” .
- [8] BRASIL, 2006a. “Manual Técnico do Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose Animal.” .
- [9] BRASIL, 2006. “Instrução Normativa 6/2004. Regulamento Técnico Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose.” .
- [10] CHARRO, F., OSÓRIO, A., 2009. “Diagnóstico da tuberculose bovina por meio de histopatologia”. Disponível em: <<http://www.propp.ufms.br/gestor/titan.php?target=openFile&fileId=429>>. Acesso em 18 de Agosto de 2015.
- [11] CLARKE, C., ET AL., 2001. “Manual Merck de Veterinária” . .

- [12] COBOS-MARIN, L., MONTES-VARGAS, T., RIVIERA-GUTEIERREZ, S., et al., 2003, “A novel multiplex-PCR for the rapid identification of *Mycobacterium bovis* in clinical isolates of both veterinary and human origin”, *Epidemiology and Infection*, v. 130, pp. 485–490.
- [13] CONLAN, A., POLLOCK, E., MCKINLEY, T., et al., 2015, “Potential Benefits of Cattle Vaccination as a Supplementary Control for Bovine Tuberculosis”, *Public Library of Science computational biology*, v. 11, n. 2.
- [14] CONOVER, W., 1980, *Practical Nonparametric Statistics*, v. 2. New York, John Wiley & Sons, Inc.
- [15] CORNER, L., 2006, “The role of wild animal populations in the epidemiology of tuberculosis in domestic animals: How to assess the risk”, *Vet Microbiol.*, v. 112, pp. 303–312.
- [16] CORNER, L., 1994, “Post mortem diagnosis of *Mycobacterium bovis* infection in cattle.” *Veterinary Microbiology*, v. 40, pp. 53–63.
- [17] CORRÊA, W., CORRÊA, C., 1992, *Enfermidades infecciosas dos mamíferos domésticos*, v. 843. Rio de Janeiro, MEDSE.
- [18] COSTA, L., 2012. “Caracterização da tuberculose bovina em regiões de relevância econômica no Estado da Bahia.” .
- [19] COURCOUL, A., MOYEN, J., BRUGERE, L., et al., 2014, “Estimation of sensitivity and specificity of bacteriology, histopathology and PCR for the confirmatory diagnosis of bovine tuberculosis using latent class analysis”, *PloS one*, v. 9.
- [20] DINIZ, C., LOUZADA-NETO, F., 2000. “Data Mining: uma introdução” . .
- [21] DUDA, R., HART, P., STORK, D., 2001. “Pattern classification” . .
- [22] FIGUEIREDO, E., SILVESTRE, F., CAMPOS, W., et al., 2008, “identification of *Mycobacterium bovis* isolates by multiplex PCR”, *Brazilian Journal of Microbiology*, v. 39, pp. 1–4.
- [23] FIGUEIREDO, E., SILVESTRE, F., CAMPOS, W., et al., 2009, “Identification of *Mycobacterium bovis* Isolates by a multiplex PCR”, *Brazilian Journal of Microbiology*, v. 40, pp. 231–233.
- [24] FLETCHER, R., FLETCHER, S., WAGNER, E., 1991. “Epidemiologia Clínica.” .

- [25] FRÁGUAS, S., CUNHA-ABREU, M., FERREIRA, A., et al., 2008, “Estudo comparativo de métodos complementares para o diagnóstico da tuberculose bovina em animais reagentes a tuberculinização”, *Revista Brasileira de Ciência Veterinária*, v. 15, pp. 117–121.
- [26] FRIEDMAN, H., PAPPER, S., 1985. “Manual do diagnóstico em medicina interna.” .
- [27] FURLANETTO, L. V., FIGUEIREDO, E., CONTE JÚNIOR, C., et al., 2012, “Uso de métodos complementares na inspeção post mortem de carcaças com suspeita de tuberculose bovina”, *Pesquisa Veterinária Brasileira*, v. 32 (11), pp. 1138–1144.
- [28] GRACIOSO, A., SOUZA FILHO, R., GONZAGA, A., et al., 2013, “A computer software system to assist in the diagnosis of animal tuberculosis”, *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, v. 65, n. 2, pp. 622–626.
- [29] GUTIERREZ, M., TELLECHEA, J., GARCIA MARIN, J., 1998, “Evaluation of cellular and serological diagnostic tests for the detection of Mycobacterium bovis-infected goats”, *Vet Microbiol.*, v. 62 (4), pp. 281–290.
- [30] HAAGSMA, J., 1995. “Bovine tuberculosis” . .
- [31] HAYKIN, S., 2008, *Redes Neurais: Princípio e prática*, v. 2. Porto Alegre, RS, Bookman.
- [32] JOHNSON, R., WICHERN, D., 1999, *Applied multivariate statistical analysis*. New Jersey, Prentice-Hall.
- [33] JONES, T., HUNT, R., KING, N., 2000, *Patologia veterinária*, v. 6. Brasil, Manole.
- [34] JORGE, M., GOTLIEB, S., LAURENTI, R., 2001. “A saúde no Brasil: análise do período 1996-1999.” .
- [35] KANTOR, I., RITACCO, V., USABIAGA, J., 1994, “Bovine tuberculosis in Latin America and Caribbean: current status, control and eradication programs”, *Vet Microbiol.*, v. 40, pp. 5–14.
- [36] KHANNA, T., 1990. “Foundation of Neural Networks.” .
- [37] LILENBAUM, W., 2000, “Atualização em Tuberculose Bovina. Uma minirevisão.” *Rev. Bras. Med. Vet.*, v. 22, n. 4, pp. 145–151.



- [38] LOBUE, P., BETACOURT, W., PETER, C., et al., 2003, “Epidemiology of Mycobacterium bovis disease in San Diego County”, *The International Journal of Tuberculosis and Lung Disease*, v. 7, pp. 180–185.
- [39] MARASSI, C. D., MEDEIROS, L., FIGUEIREDO, E., et al., 2013, “A multi-disciplinary approach to diagnose naturally occurring bovine tuberculosis in Brazil”, *Pesquisa Veterinária Brasileira*, v. 33 (01), pp. 15–20.
- [40] MARASSI, C., ZARDEN, C., OELEMANN, W., et al., 2014, “Detection of specific antibodies in cows after injection of PPD”, *Pesquisa Veterinária Brasileira*, v. 34 (03), pp. 241–243.
- [41] MCCULLOCH, W., PITTS, W., 1943, “A logical calculus of the ideas immanent in nervous activity.” *Mathematical Biophysics*, v. 5, pp. 18–27.
- [42] MEHTA, J., SRINIVASAN, R., 1970, “Behrens-Fisher problem.” *Biometrika*, v. 7, pp. 649–655.
- [43] MICHEL, A., MULLER, B., VAN HELDEN, P., 2010, “Mycobacterium bovis at the animal–human interface: 13 A problem, or not?” *Veterinary Microbiology*, v. 140, pp. 371–381.
- [44] MONAGHAN, M., DOHERTY, M., COLLINS, J., et al., 1994, “The tuberculin test.” *Veterinary Microbiology*, v. 40, pp. 111–124.
- [45] MORRISON, D., 2005, *Multivariate Statistical Methods*. New York, 4. ed. McGraw–Hill.
- [46] NEILL, S., POLLOCK, J., BRYSON, D., et al., 1994, “Pathogenesis of Mycobacterium bovis infection in cattle.” *Veterinary Microbiology*, v. 40, pp. 41–52.
- [47] NORVING, P., RUSSEL, S., 2004. “Intelligence Artificial” . .
- [48] O’REILLY, L., DABORN, C., 1995, “The epidemiology of Mycobacterium bovis infections in animals and man: a review.” *Tubercle and Lung Disease*, v. 76, pp. 1–46.
- [49] PAIXÃO, T., MINHARRO, S., CARVALHO NETA, A., et al., 2008, “O diagnóstico *post mortem* da tuberculose bovina”, *Cadernos Técnicos de veterinária e Zootecnia*, v. 59, pp. 26–42.
- [50] PARRA, A., GARCÍA, N., GARCÍA, A., et al., 2008, “Development of a molecular diagnostic test applied to experimental abattoir surveillance on bovine tuberculosis”, *Veterinary Microbiology*, v. 127, pp. 315–324.

- [51] PINTO, P., VILORIA, M., FARIA, J., et al., 2004, “Avaliação do desempenho dos exames anatomopatológico e histopatológico na inspeção post mortem de bovinos suspeitos ou reagentes a prova de tuberculinização”, *Revista Brasileira de Ciência Veterinária*, v. 11, pp. 27–31.
- [52] RADOSTITS, O., BLOOD, D., GAY, C., et al., 2002, *Clínica veterinária: um tratado de doenças dos bovinos, ovinos, suínos, caprinos e equinos*, v. 1877. Rio de Janeiro, Guanabara Koogan.
- [53] RIET-CORREA, F., SCHILD, A., MENDEZ, M., et al., 2001, *Doenças de Ruminantes e Equinos.*, v. 1. São Paulo, Varela.
- [54] RODRIGUEZ, C., ZUMÁRRAGA, M., OLIVEIRA, E., et al., 2004, “Caracterização molecular de isolados de *Mycobacterium bovis* do Estado de São Paulo Brasil, utilizando a técnica de spoligotyping”, *Arquivos do Instituto Biológico*, v. 71, pp. 277–282.
- [55] ROSENBERGER, G. E. A., 1989, “Enfermidades de los Bovinos”, *Buenos Aires: Editorial Hemisferio Sul S.A.*, v. 2, pp. 139–151.
- [56] ROXO, E., 2008. “Tuberculose humana e animal”. Disponível em: <[http://www.infobibos.com/artigos/2008\\_1/tuberculose/index.htm](http://www.infobibos.com/artigos/2008_1/tuberculose/index.htm)>. Acesso em: 06.02.2015.
- [57] ROXO, E., 1996, “Tuberculose bovina: revisão.” *Arquivos do Instituto Biológico*, v. 63, pp. 91–97.
- [58] RUGGIERO, A., IKUNO, A., FERREIRA, V., et al., 2007, “Tuberculose bovina: alternativas para o diagnóstico (Bovine tuberculosis: alternatives for the diagnosis)”, *Arquivos do Instituto Biológico*, v. 74, pp. 55–65.
- [59] VORDERMEIER, M., JONES, G., WHELAN, A., 2011, “DIVA reagents for bovine tuberculosis vaccines in cattle”, *Informa Healthcare London*.
- [60] WEBB, A., 2002. “Statistical Pattern Recognition” . .
- [61] ZANINI, M., MOREIRA, E., LOPES, M., 2001, “*Mycobacterium bovis* polymerase chain reaction identification in bovine lymph node biopsies and genotyping in isolates from southeast Brazil by spoligotyping and RFLP.” *Memórias do Instituto Oswaldo Cruz*, v. 96, pp. 809–813.